Version 2018-04-08/TE

Study mortality with hazard rates, not probabilities

Torbjørn Ergon[1*], Ørnulf Borgan[2], Chloé Rebecca Nater[1] and Yngvild Vindenes[1]

[1] Centre for Ecological and Evolutionary Synthesis, Department of Biosciences, University of Oslo, P.O. Box 1066 Blindern, N-0316 Oslo, Norway

[2] Department of Mathematics, University of Oslo, P.O. Box 1066 Blindern, N-0316 Oslo, Norway

[*] Corresponding author: e-mail: t.h.ergon@ibv.uio.no, phone: (+47) 22857311 / (+47) 92602138

Running headline: Study mortality with hazard rates

Abstract

1. Mortality is a key process in ecology and evolution, and much effort is spent on statistical and theoretical modelling of this process. Mortality hazard rates describe individuals' instantaneous ability to survive at different times, whereas survival probabilities are defined for time-intervals of a given length. In this commentary, we argue that it is often more meaningful to model and interpret interval-specific time-averaged mortality hazard rates than survival probabilities, also when using discrete-time models.
2. We discuss three topics where the concept of hazard rates is essential for sound biological inference, but nevertheless often not used: (i) modelling of covariate effects on survival probabilities or mortality hazard rates, (ii) modelling of multiple sources of mortality and competing risks, and (iii) elasticity analyses of population growth rate to demographic parameters. To facilitate estimation of cause-specific mortality hazard rates, we provide amendments to R package 'marked'.
3. By focusing on modelling mortality hazard rates, one avoids several inherent problems of comparing survival or mortality probabilities. In particular, interpretations about relative differences between mortality hazard rates (measurements on a ratio scale), or effects of relative changes in mortality hazard rates on population growth (elasticities), are more meaningful than similar interpretations involving survival (or mortality) probabilities or odds (measurements on an absolute scale).
4. The concept of hazard rates is essential for understanding ecological and evolutionary processes and we give an intuitive explanation for this, using several examples. We provide some practical guidelines and suggestions for further methods developments.

Keywords: Capture-mark-recapture, cause-specific mortality, competing risks, sensitivity/elasticity analysis, measurement theory, time-averaged hazard rates, loglog link function, odds-ratio vs. hazard-ratio.

Introduction

Mortality takes place in continuous time, but statistical and theoretical models in ecology are often formulated on a discrete time-scale, making it necessary to represent the mortality process with **survival probabilities** and cause-specific **mortality probabilities** (terms in bold at first occurrence are defined in the Glossary). While **mortality hazard rate** ($h(t)$) is the instantaneous intensity of deadly events that an individual is exposed to (Box 1), the probability of surviving a given interval from time t_1 to time t_2 ($S_{t_1 \to t_2}$) is a function of the integrated hazard rate over the interval,

$$S_{t_1 \to t_2} = e^{-\int_{t_1}^{t_2} h(t)dt} \tag{1}$$

(Box 1). The integral in this equation can be replaced by the average hazard rate within the interval times the length of the interval such that survival probability in interval j becomes

$$S_j = e^{-\bar{h}_j \Delta_j} \tag{2}$$

where Δ_j is the length of the time-interval (usually one time-unit) and the unit of measurement for the time-averaged hazard rate, $\bar{h}_j$, is number of deadly events per time-unit (Collett 2014). Such interval-specific survival probabilities and time-averaged hazard rates are often sufficient for specific analyses where fluctuations within intervals (e.g. due to seasonal or daily cycles) are not relevant.

Although survival probabilities are simple transformations of time-averaged mortality hazard rates (eq. (2)), there are some important differences in how these two measurements can be interpreted and modelled. The survival probability is a unit-free measurement on an **absolute scale**, meaning that no transformation can be performed without changing its *meaning* (Houle *et al.* 2011). For example, if S is the probability of surviving a whole year, then $S^{1/2}$ can be interpreted as the probability of surviving half the year, assuming that survival probabilities in the first and second half of the year are equal. The mortality hazard rate, on the other hand, is a **ratio-scaled measurement**, and its unit of measurement may thus be changed (e.g., from 'year^{-1}' to 'month^{-1}') by simply multiplying with a positive number (just like a given distance can be measured in 'meters' or 'feet') (Houle *et al.* 2011).

Any model based on survival probabilities can be rewritten in terms of time-averaged hazard rates using eq. (2). The relationship between cause-specific mortality probabilities and hazard rates is

somewhat more complex, and we will return to this later. In the fields of statistical modelling of capture-recapture data and structured population modelling, there is a tradition for basing inference on (differences in) survival probabilities (Caswell 2001; McCrea & Morgan 2015). In other fields, such as in fisheries, medicine and human demography, it is more common to model and interpret mortality hazard rates (Quinn & Deriso 1999; Keyfitz & Caswell 2005; Collett 2014). Sometimes, this choice of approaches has important inferential consequences, and we call for more attention to these issues in ecological modelling. In general, modelling based on mortality hazard rates provides a better basis for biological inference and avoids several inherent problems of comparing survival or mortality probabilities. Our aim is here to explain and illustrate this claim.

In the first section below, we will argue that effects of covariates on survival probabilities are better described by log-hazard (loglog-link) models instead of the commonly used log-odds (logit-link) models. Under the second heading, we discuss and illustrate the use of hazard rate models to model multiple competing sources of mortality. In the third section, we question the tradition of basing inferences on sensitivities and elasticities of population growth rate or fitness to survival probabilities rather than mortality hazard rates. Finally, we conclude with some perspectives on future developments. All sections are accompanied by further details, including R-code, in the online Supporting Information.

Logit-linear models of survival vs. log-linear models of mortality hazard rates

Understanding effects of different covariates (e.g. relating to climate or individual properties) on survival/mortality is a central aim in ecology and evolution. Such effects are typically estimated using linear models of survival probabilities transformed by a link-function (Lebreton *et al.* 1992). When a **logit-link** function is used for survival probabilities, the exponents of **linear contrasts** may be interpreted as **odds-ratios**, whereas, when a **loglog-link** is used, the exponents of linear contrasts become **hazard ratios** (Box 1). In other words, when using a logit-link, the **relative difference** in survival **odds** is modelled, while using a loglog-link implies relative effects on mortality hazard rates (Box 1 and Appendix S1 in the Supporting Information).

While most practitioners are presumably aware of the different functional forms implied by the choice of link-function when using continuous covariates, there seems to be less awareness of the fact that parameterising the models with survival probabilities representing different durations (e.g. months instead of years) also changes the functional forms when using any other link-function than the loglog-link (Figure 1; Appendix S1 includes R code to reproduce all figures in the article and provides additional explanations). That is, when *not* using a loglog-link for survival probabilities, selecting the time-unit for the analysis (which can be set in commonly used computer programs such as MARK (White & Burnham 1999)) is part of the model definition. In many studies, it is natural to model

survival on an annual time scale as this represents survival over a full seasonal cycle. However, for species with a life-expectancy of much less than a year, such as many small rodents, yearly survival is not meaningful, and one may choose, somewhat arbitrarily, to model weekly or monthly survival instead. Sometimes it may also be of interest to estimate the effects of certain covariates on survival until a specific life-stage (e.g. the effect of chick body mass on survival until reproduction in a long-lived seabird). If a logit-link is used in such cases, one will get different results depending on the choice of time unit that the survival probabilities in the model represent. Using a loglog-link for survival, which implies a log-linear model for the time-averaged mortality hazard rate (Box 1), avoids the problem of selecting the time-spans that the survival probabilities represent because survival over any given interval is then invariant to the interval length used in the modelling of covariate effects (Figure 1; Appendix S1). Further, mortality hazard rates provide a cohesive basis for fitting hazard rate functions of age (Gaillard *et al.* 2004), to model seasonal changes in mortality as continuous functions of time (Ergon, Yoccoz & Nichols 2009), to allow transitions between individual states in continuous time (Miller & Andersen 2008; Ergon, Yoccoz & Nichols 2009; Conn, Cooch & Caley 2012; Choquet *et al.* 2017), or to accommodate analysis of data where individuals are not encountered during discrete occasions (Choquet *et al.* 2017).

While survival probabilities usually provide meaningful inference, differences in survival expressed as odds-ratios are often difficult to interpret in a biological context. For example, consider an environment where yearly juvenile survival is 0.25 (odds = 0.25/(1-0.25) = 1/3) and yearly adult survival is 0.50 (odds = 1). Then the odds of surviving one year is 3 times higher for adults compared to juveniles, but the odds of surviving two years (which may correspond to the age at maturation) is 5 times higher for adults than juveniles (1/3 vs. 1/15). Now consider an improved environment where yearly survival of juveniles and adults are respectively 0.50 and 0.71, the odds ratios for surviving one year is 2.4 (and 3.0 for surviving two years), suggesting a smaller difference between juveniles and adults than in the previous environment. In both environments, however, the time-averaged mortality hazard rate (eq. 2) for juveniles is 2 times higher than the mortality hazard rate for adults, meaning that the intensity at which deadly events occur is twice as high for juveniles compared to adults, irrespective of the environment and the time-scale used. Hence, as illustrated in Figure 2, an odds-ratio for any given hazard ratio does not only depend on the length of the time-interval that the odds apply to, but also on the survival probability of the reference group. In contrast, hazard ratios have a clear biological interpretation, and facilitate meaningful comparisons of effect sizes (Box 1).

Modelling of multiple sources of mortality and competing risks

Discerning between different sources of mortality is importance key challenge in ecology, evolutionary biology and population management (Schaub & Lebreton 2004; Péron 2013; Wolfe *et al.* 2015). To understand selective forces and population dynamics, it is also important to examine how

different sources of mortality co-vary and interact (Nichols *et al.* 1984; Boyce, Sinclair & White 1999; Koons *et al.* 2014). For example, an essential question is whether increased hunting or harvesting will lead to a reduction (or potentially an increase) in mortality due to other causes. Such effects can result from density dependent responses or non-random removal of weaker (or stronger) individuals (i.e., selection) in a heterogeneous population. Changes in one source of mortality may also lead to evolutionary changes in other mortality rates or life-history traits. For example, age-dependent mortality patterns affect optimal investments in development and maintenance/longevity as well as reproductive effort at different ages (Charlesworth 1980; Caswell 2007a; Koons *et al.* 2014).

When studying multiple sources of mortality, it is essential to realise that if the hazard rate representing one cause of mortality (say, hunting) increases, the probability of dying from any other cause will automatically be reduced, even if the mortality hazard rates representing the other causes remain unchanged. This is simply because it has now become more likely that the individual will die of this cause (hunting) before other deadly events can occur (Appendix S2 in the Supporting Information). We may say that the different causes of mortality "compete" over killing the individual first, and the probability of "winning" depends on how "strong" the competitors are (referred to as "competing risks" in survival analysis (e.g. Collett 2014)).

If we assume that there are K causes of mortality and the mortality hazard rates for each cause k, $h_k(t)$, remain proportional within a time interval, such that the fraction $\alpha_k = h_k(t)/\sum_{l=1}^{K} h_l(t)$ remains constant for all causes throughout the interval, it can be shown (see Appendix S2) that the probability of dying from cause k (i.e., the '**risk**' of cause k) is

$$P_k = (1 - S)\alpha_k \,, \tag{3}$$

where S is the probability of surviving the interval (from t_1 to t_2) defined to have length 1. That is,

$$S = e^{-\int_{t_1}^{t_2} \sum_{k=1}^{K} h_k(t)dt} = e^{-\sum_{k=1}^{K} \int_{t_1}^{t_2} h_k(t)dt} = e^{-\sum_{k=1}^{K} \bar{h}_k} \,, \tag{4}$$

where $\bar{h}_k$ is the time-averaged hazard rate for cause k (eq. (2)). In the most general model these parameters would be interval- and individual-specific, but indices for interval and individual are omitted here for simplicity. Note that when the hazard rates $h_k(t)$ remain proportional within an interval, the relative differences between them are the same as the relative differences between the time-averaged hazard rates. Hence, we can also interpret α_k as

$$\alpha_k = \frac{\bar{h}_k}{\sum_{l=1}^{K} \bar{h}_l}\,. \tag{5}$$

When information about cause of death (e.g., harvesting) is available for at least a subset of marked individuals in the population, multi-state capture-recapture models can be fitted to the data to address questions relating to e.g. compensatory effects of harvesting or age-related patterns in mortality from different causes (Servanty *et al.* 2010; Koons *et al.* 2014). Such models involve probabilities of remaining in an "alive-state" (survival probability S) or transitioning to any of the absorbing death-by-cause states (probabilities P_k's). From equations (3), (4) and (5) we see that there are at least three distinct ways to parameterize these transition probabilities when hazard rates within intervals are assumed to be proportional: [1] with S and P_k's under the constraint that $S + \sum_{k=1}^{K} P_k = 1$; [2] with S and α_k's under the constraint that $\sum_{k=1}^{K} \alpha_k = 1$; or [3] with the cause-specific time-averaged mortality hazard rates $\bar{h}_k$'s. The relationships between the parameters in these three parameterizations are given in Table 1. Note that when constraining these individual- and interval-specific parameters by functions of individual age or covariates, these alternative model formulations will generally lead to different models.

Since a cause-specific mortality probability P_k does not only depend on the mortality hazard rate for cause k, but also on the total mortality hazard rate (Appendix S2), it will in general be biologically more interesting to model and interpret the cause-specific mortality hazard rates (h_k's). Nevertheless, statistical modelling of capture recapture data incorporating multiple (usually two) sources of mortality, has so far, to our knowledge, only been based on parameterizations for transition probabilities that do not include hazard rates explicitly (first two parameterizations in Table 1). Commonly used software for fitting multi-state capture-recapture data such as MARK (White & Burnham 1999) and E-SURGE (Choquet, Rouan & Pradel 2009) use the multinomial logit-link function to constrain the multinomial probability vector to sum to 1, which does not facilitate modelling of cause-specific hazard rates (Appendix S2). In the supplementary material we provide an R function that can be used with the 'marked' package (Laake, Johnson & Conn 2013) to model cause-specific hazard rates in capture-recapture analyses and give an example using simulated data in Appendix S3. Even though methods for estimating cause-specific mortality hazard rates from time-to-event data have been more easily available (Lunn & McNeil 1995), the majority of ecological studies based on such data from telemetry or loggers also put undue emphasis on mortality probabilities rather than hazard rates.

Many studies have estimated relationships between probabilities of dying from different causes across years or space in order to assess whether mortality from these causes are additive (not

associated), compensatory (negatively associated) or depensatory (positively associated) (Bischof *et al.* 2009; Servanty *et al.* 2010; Péron 2013; Wolfe *et al.* 2015). One problem of using mortality probabilities in such studies is that they are, as explained above, intrinsically related; increasing the mortality hazard rate of one cause implies that the probability of dying from this mortality cause will increase, while the probability of dying from all other causes decreases (Appendix S2). As shown in Figure 3, this intrinsic negative association may dominate empirical correlations between cause-specific mortality probabilities across years (or groups of individuals) even when correlations between the cause-specific mortality hazard rates across years (or groups) are strongly positive. Realizing this problem, researchers have come up with innovative ways to reduce "bias"[1] in correlations between cause-specific mortality probabilities (Schaub & Lebreton 2004; Servanty *et al.* 2010) or to interpret these correlations with respect to hypotheses for compensatory versus additive mortality patterns (Péron 2013). However, when modelling mortality hazard rates, no such "corrections" are necessary to obtain the desired interpretation; when partitioning mortality into one specific cause (e.g. "harvesting/predation") and all other causes, a negative correlation between the mortality hazard rates among years indicates compensation, a positive correlation indicates depensation, and a correlation close to zero may indicate additivity. Furthermore, a near constant sum of the hazard rates indicates perfect compensation, while a negative correlation between the sum of the hazard rates and the harvesting hazard rate indicates overcompensation. When assessing such correlations it is important to ensure that the process correlations are not confounded with sampling correlations, for example by using appropriate hierarchical models (Royle & Dorazio 2008).

Our recommendation to model and interpret mortality hazard rates instead of mortality probabilities is also highly relevant for studies of age related patterns in mortality. For instance, Koons *et al.* (2014) assessed patterns of senescence in a population of snow geese (*Chen caerulescens*) and a population of roe deer (*Capreolus capreolus*) by estimating age-dependent cause-specific mortality probabilities pertaining to hunting/human-related causes as well as natural causes. For both species they concluded that the yearly probability of dying from hunting/human-related causes was nearly independent of age, while the probability of dying from natural causes increased sharply at higher ages. However, as explained above, for the probability of dying from human-related causes to remain constant while the probability of dying from other causes increases with age, the mortality hazard rate representing human related causes must also increase to compensate for the loss of opportunities (Figure 4).

[1] Note that correlations between mortality probabilities only are "biased" in the sense they do not represent correlations between cause-specific mortality hazard rates. They are not biased when interpreted strictly as correlations between interval-specific mortality probabilities.

Elasticity analyses based on mortality hazard rates instead of survival probabilities

Sensitivity analyses are essential in studies of population dynamics, management of natural populations, life-history theory and evolutionary modelling (Benton & Grant 1999; van Tienderen 2000; Caswell 2001; Caswell 2007b). In particular, many studies consider the sensitivity of the asymptotic population growth rate (λ) to changes in demographic parameters such as survival probability (S) and fecundity (F), as well as lower-order parameters and environmental variables. It is common to interpret and compare such sensitivities with respect to different demographic parameters on a relative (logarithmic) scale. That is, the relative change in growth rate per relative change in, for example, survival during one life-stage, referred to as the *elasticity* of λ to survival $\left(\epsilon_S = \frac{\partial \log \lambda}{\partial \log S}\right)$, is compared to the elasticity of λ to fecundity (Caswell 2001).

One motivation for using elasticities instead of sensitivity analysis of **absolute effects** to absolute changes, is that a unit change in transition probabilities such as survival probabilities is not comparable with a unit change in fecundity (Caswell 2001, p. 226; Link & Doherty 2002). However, it can be argued that also a comparison between the effects of relative changes (elasticities) in a variable of an absolute measurement scale type (e.g. survival probabilities) and a ratio scaled variable (e.g. fecundity) is not meaningful (Houle *et al.* 2011). It seems much more meaningful to compare elasticities for two ratio-scaled variables such as mortality hazard rate and fecundity. Moreover, a certain environmental change, management action or genetic change is not likely to have the same relative effect on a high vs. low survival probability (e.g., a 10% increase in survival could represent a change from 0.10 to 0.11 or from 0.90 to 0.99, and a 10% increase is impossible when survival probability is 0.95). Instead, it seems generally more plausible that such effects will have a proportional effect on the intensity at which deadly events occur; i.e., the mortality hazard rate (see Box 1).

The difficulties in biological interpretation of elasticities with respect to survival probabilities are also made clear by considering elasticities of some transformations of survival probabilities. Link and Doherty (2002) pointed out that elasticity of λ to survival probability (S) is only equal to the negative elasticity of λ to probability of mortality ($1 - S$) when $S = 0.5$. These two elasticity measures represent very different degrees of change in the same underlying mortality process when S is not close to 0.5, but it is hard to justify why any of them should be more biologically relevant than the other (Link & Doherty 2002). Further, elasticity to a survival probability depends critically on the length of the time period that it applies to. For example, instead of calculating elasticities to yearly survival probability (S_y) during the pre-reproductive life-stage, one could calculate elasticities to survival over two years (which may correspond to age at maturation). If we, for simplicity, assume that survival probabilities through the first and second year are the same, survival over the two years

becomes $S_{2y} = S_y^2$. Population growth rate is naturally more sensitive to a relative increase in S_y than to the same relative increase in S_{2y}; elasticity of λ to S_y is twice the elasticity of λ to S_{2y} (Appendix S1). Now consider instead the elasticity of λ to the time-averaged mortality hazard rate, $\bar{h} = -\log(S)$, which is $\epsilon_{\bar{h}} = \log S \frac{\partial \log \lambda}{\partial \log S} = -\bar{h}\epsilon_S$. This elasticity is invariant to any re-parameterizations on the form $S_{2y} = S_y^2$ because $\log S_{2y} \frac{\partial \log \lambda}{\partial \log S_{2y}} = \log S_y \frac{\partial \log \lambda}{\partial \log S_y}$ (Appendix S1). Note also that the negative elasticity to mortality hazard rate is the elasticity to "current life-expectancy" (Appendix S1). Figure 5 shows the relationships between elasticities of λ to survival probability and mortality hazard rate when the model is parameterized with survival probabilities over different time intervals. Clearly, if the purpose of the elasticity analysis is to investigate the effect of a relative change in the ability to survive during an interval of a somewhat arbitrary length, it is more appropriate to base the elasticity analysis on the time-averaged mortality hazard rate ($\bar{h}$) than on the probability of surviving this interval (S).

In the example used in Figure 5a, where yearly juvenile survival probability is much lower than yearly adult survival probability, population growth rate has a higher elasticity to adult survival than to juvenile survival. This pattern is typical of long-lived vertebrates, and is commonly interpreted as population growth rate or fitness being more sensitive to survival of adults than juveniles in such populations (see e.g. references in Gaillard and Yoccoz (2003)). However, elasticity with respect to mortality hazard rates shows the *opposite* pattern; elasticity to juvenile mortality hazard rate is higher than to adult mortality hazard rate. These counterintuitive differences do not occur when juvenile survival and fecundity are sufficiently high compared to adult survival (Figure 5b). Figure 6a shows how elasticities of λ with respect to survival probabilities relate to juvenile and adult survival, and Figure 6b shows the equivalent relationship for elasticities of λ with respect to mortality hazard rates. These very different patterns are due to the fact that a relative change in mortality hazard rate leads to a small relative change in survival probability when survival is high, but a much larger relative change when survival probability is low (Figure 7).

Sensitivity/elasticity analysis also plays an important role in our understanding of underlying mechanisms to population growth and extinction risk in fluctuating environments. Several studies have compared sensitivities/elasticities of the average population growth rate (or fitness) to demographic parameters in relation to their variability over time (reviewed in McDonald *et al.* (2017)). Most analyses conclude that life-history traits to which fitness is most sensitive tend to be more buffered (i.e., canalised) against environmental fluctuations than other traits. However, some authors have pointed out that such conclusions may be highly dependent on the measurement scale of the life-history traits (references in McDonald *et al.* (2017)). In particular, there has been a concern that among-year variance in survival and state transition probabilities are bounded by a maximum attainable variance given by their mean, and hence cannot be as variable as other traits. Using variance-stabilising transformations as advocated by Link and Doherty (2002), McDonald *et al.* (2017)

reported on an analysis of demographic strategies in a wide range of plants and found that logit-transformed survival probabilities were in fact often highly labile (i.e. not canalised) despite strong variance-standardised sensitivities (VVS_q) for these traits (VVS_q is here defined as $VVS_q(\lambda, S) = \text{Sensitivity}(\log(\lambda), q(S))$ where $q(S)$ is variance-standardised transformation of survival; Link and Doherty 2002). The use of the logit function as a variance-stabilising transformation of survival probabilities in this analysis assumes that among-year variance in survival probability is proportional to $\mu_S^2(1-\mu_S)^2$ where μ_S is the mean survival among years. Alternatively, the same analysis could have used a loglog-transformation of survival probabilities which would stabilize the variance when the variance in survival probability is proportional to ${\mu_S}^2 \log(\mu_S)^2$ (Appendix S1). Since among-year variability in survival probability is presumably highly system specific as it depends on complex trophic interactions and (non-linear) environmental effects relating to for example predators, pathogens, climate and intraspecific population density, it is difficult to argue that one variance-mean relationship is generally more applicable than others. Hence, we do not advocate the use of loglog-transformation of survival probabilities due to its particular variance-stabilizing properties, but instead because this transformation is equivalent to a log-transformation of mortality hazard rates (Box 1). This means that the variance-stabilized sensitivity for survival ($VVS_q(\lambda, S)$) using a loglog-transformation ($q(S) = \log(-\log(S))$) is identical to the elasticity for mortality hazard rate ($VVS_q(\lambda, S) = \epsilon_{\bar{h}}$) (Appendix S1). As argued above, it is more meaningful to compare elasticities of two ratio-scaled variables such as mortality hazard rate and fecundity than relying on some rather arbitrary transformations of survival probabilities for such comparisons.

Conclusion and perspectives

Mortality is a process in continuous time, and mortality hazard rates are the most fundamental measurements describing this process. We have here highlighted the importance of basing inference on hazard rates also in discrete-time models. In models where all transition probabilities are survival probabilities, we recommend modelling log-linear (multiplicative) effects on time-averaged hazard rates by using a loglog-link function for survival probabilities. In models involving multiple causes of death, estimates of hazard rates can be obtained by using a hazard rate based multinomial link function, assuming that the cause-specific hazard rates remain proportional within intervals (Table 1). This allows direct inference regarding interactions between different sources of mortality. When the assumption about proportionality within intervals is not reasonable one may split up the intervals such that e.g. hunting mortality can only occur during the hunting season.

The issues we have discussed here apply more generally to any process involving transitions that can take place instantaneously in continuous time, such as transitions between disease states (Conn, Cooch & Caley 2012) or reproductive states of multivoltine species (Ergon, Yoccoz & Nichols 2009). When more than one transition can take place at any point in time within the intervals, an

increase in the hazard rate for one transition will reduce the probability that the other competing transitions will occur. As a general method, transition probability matrices may be computed from matrices of transition hazard rates by the use of matrix exponentials (Keyfitz & Caswell 2005 chap. 17.5; Miller & Andersen 2008; Conn, Cooch & Caley 2012). These models may in principle include intermediate and reversible transitions, but identifiability of statistical models may be problematic. It is also often desirable to focus the modelling on the latent probability density function for time/age of state-transition rather than the corresponding hazard rate (Ergon, Yoccoz & Nichols 2009; de Valpine *et al.* 2014; Appendix S1 1.4.1.4 and 1.4.1.5).

We conclude that large parts of the ecological literature put an undue emphasis on modelling and interpreting state transition probabilities (such as survival and mortality probabilities), when inference based on (time-averaged) transition hazard rates, or their corresponding density functions for time of transition, would be more appropriate. Methods focusing on the modelling of hazard rates are well established and much used in other fields, but are to a large extent lacking from ecologists' toolbox. This is probably due to insufficient understanding and a lack of attention to meaningful inference, as well as tradition and the fact that hazard rate based models (other than loglog-link models of survival probabilities) have not yet been implemented in software commonly used by ecologists.

Acknowledgements

We greatly appreciate comments by Guillaume Péron, Nigel G. Yoccoz and one anonymous reviewer on earlier versions of the manuscript.

Authors' contributions

TE conceived the main ideas and wrote the first version of the manuscript. All authors contributed with discussions, literature review, commenting and editing of subsequent versions.

Data accessibility

This article does not use collected data. R code for producing all figures and other supporting material is provided in the online version of the article.

Box 1 – Mortality hazard rates and survival probabilities

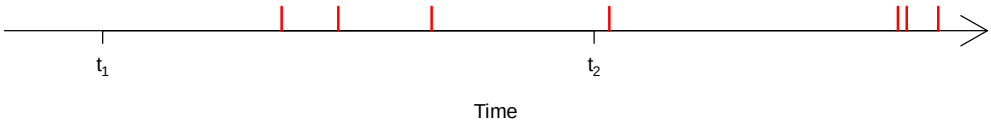


Figure B1. Realized times of recurrent events (or "deadly events experienced by an immortal individual")

A mortality hazard rate is the underlying intensity of deadly events that an individual is exposed to. To better understand the relationship between survival probabilities and hazard rates, first consider a point process of non-lethal events (e.g., ingestion of food items). An example of realized times of such events for an individual is shown in Figure B1. Denoting the intensity of this point process as a function of time as $\lambda(t)$, the expected number of events occurring during an interval from time t_1 to time t_2 is the integrated intensity over this interval, $\int_{t_1}^{t_2} \lambda(t)dt$. When events occur independently of each other, the number of events X occurring in the interval is a random Poisson variable, $X \sim \text{Poisson}\left(\int_{t_1}^{t_2} \lambda(t)dt\right)$. If these events are lethal, X must be zero for the individual to survive the interval, which according to the Poisson distribution occurs with probability $S_{t_1 \to t_2} = \Pr(X = 0) = e^{-\int_{t_1}^{t_2} \lambda(t)dt}$. When studying lethal events, the event intensity is called a 'hazard rate' and is often denoted as $h(t)$ ($= \lambda(t)$ above). A calculus-based derivation of the relationship between hazard rates and survival is provided in Appendix S1.

Mortality hazard rates may thus be interpreted as the "expected number of deadly events occurring per time-unit for an immortal individual". Since hazard rates are **ratio scaled** intensities, they may be expressed with an arbitrary time-unit and values are proportional to the length of the time-unit used (e.g, a hazard rate of one event per month is equivalent to 12 events per year). The ratio of two hazard rates, a **hazard ratio**, is thus invariant to the time-unit used to express the rates and has a clear interpretation; if the mortality hazard rate for juveniles is twice the mortality hazard rate of adults, we know that juveniles are exposed to twice as many expected deadly events per time-unit than adults. Such hazard ratios may be estimated by modelling survival probability through a **loglog-link**; $S = \exp(-\exp(\eta))$, where η is the linear predictor and $\exp(\eta)$ is the time-averaged mortality hazard rate over the interval, expressed with the interval length as a time-unit. If survival probability of one group of individuals is represented by $\eta_1 = \beta_0 + \beta_2 \log(x)$ and another by $\eta_2 = \beta_0 + \beta_1 + \beta_2 \log(x)$,

where x is a continuous covariate, the hazard ratio representing their **relative difference** in mortality, irrespective of the covariate, becomes $\frac{\exp(\eta_2)}{\exp(\eta_1)} = \exp(\beta_1)$. Further, when the covariate is multiplied by a factor C, the mortality hazard rate will be multiplied by a factor C^{β_2}(when $\beta_2 = 1$ the hazard rate is proportional to the covariate). Hence, when using a loglog-link on survival probabilities we model relative differences in hazard rates (i.e., hazard ratios). [Note that some computer programs like MARK (White & Burnham 1999) define the loglog-link function as $S = \exp(-\exp(-\eta))$, in which case the above hazard ratio becomes $\exp(-\beta_1)$.]

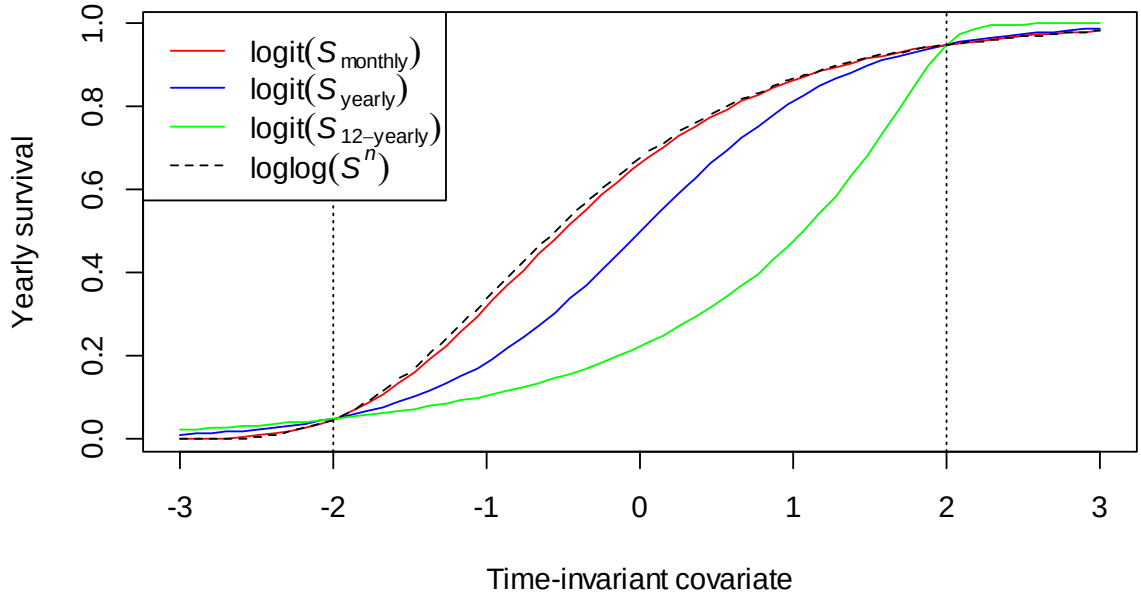


Figure 1. Functional shape of the relationship between yearly survival and a time-invariant covariate (e.g. individual specific) when using the logit link-function on survival probability over different time interval lengths (solid coloured lines; see legend). When the logit-link is applied to monthly survival (red line), it is assumed that probability of survival is the same in all of the 12 months of the year, such that $S_{\text{yearly}} = S_{\text{monthly}}^{12}$. Similarly, when the logit-link is applied to survival over 12 years, $S_{\text{yearly}} = S_{12-\text{yearly}}^{1/12}$ (green line). To facilitate comparison of functional shapes, slope and intercept parameters are defined such that yearly survival probability is 0.05 at covariate value -2 and 0.95 at covariate value 2. The dashed line shows the loglog-linear relationship between yearly survival and the covariate, which is invariant to the interval length for the survival probability. When the logit-linear relationship applies to survival probabilities over short intervals, it approaches the loglog-linear functional form (see red line).

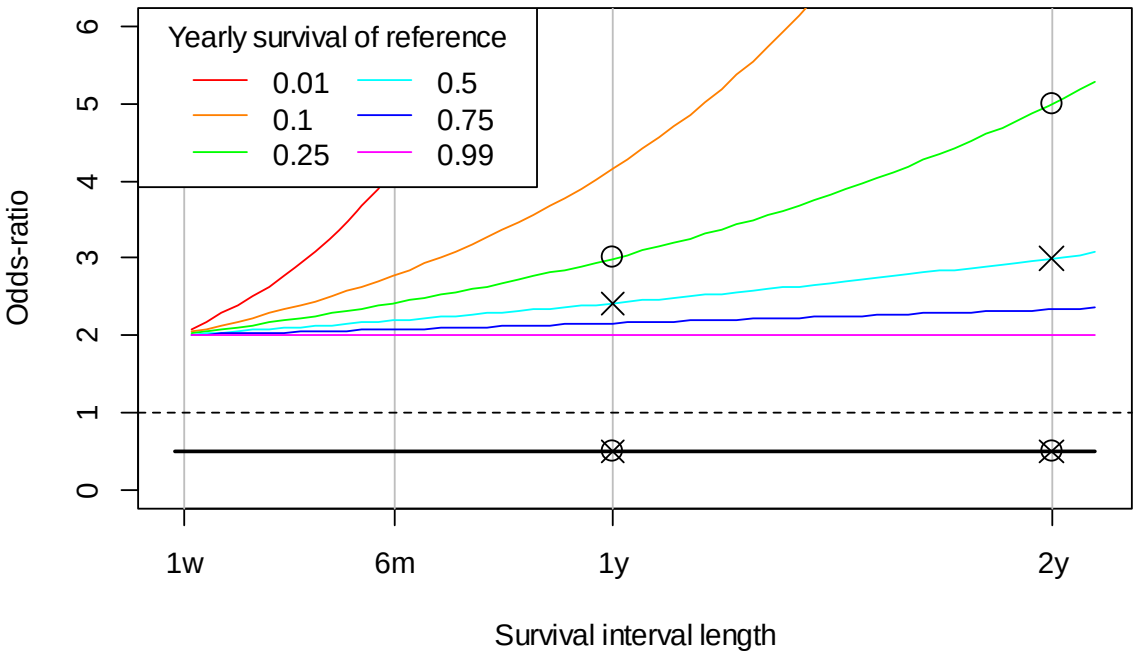


Figure 2. Contrasts between two survival probabilities expressed as odds-ratios as a function of time interval length for survival (x-axis), when the mortality hazard ratio is 0.5 (indicated with a thick black line) and yearly survival probability of the reference varies according to the colour code in the legend. Plotted points refer to the example in the text (circles represent the original environment, while crosses represent the improved environment). Vertical lines are plotted at 1 week, 6 months, 1 year and 2 years.

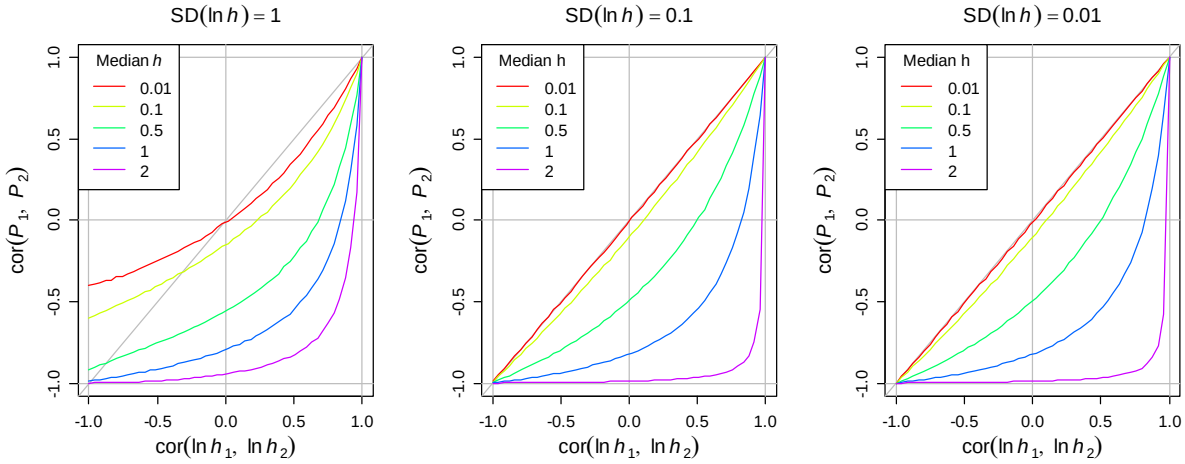


Figure 3. Correlations between probabilities of two mutually exclusive and collectively exhaustive causes of mortality (e.g. among years) as a function of the underlying correlation between the corresponding mortality hazard rates on a logarithmic scale (x-axes), assuming that the two hazard rates h_1 and h_2 have a bivariate log-normal distribution. Correlations were computed by first drawing 10^5 values of $\ln(h_1)$ and $\ln(h_2)$ and then computing the Pearson correlation coefficients of P_1 and P_2 (Table 1). The standard deviation of $\ln(h_1)$ and $\ln(h_2)$ varies by panel as indicated above the plots. Different lines represent different median hazard rates according to colour legend.

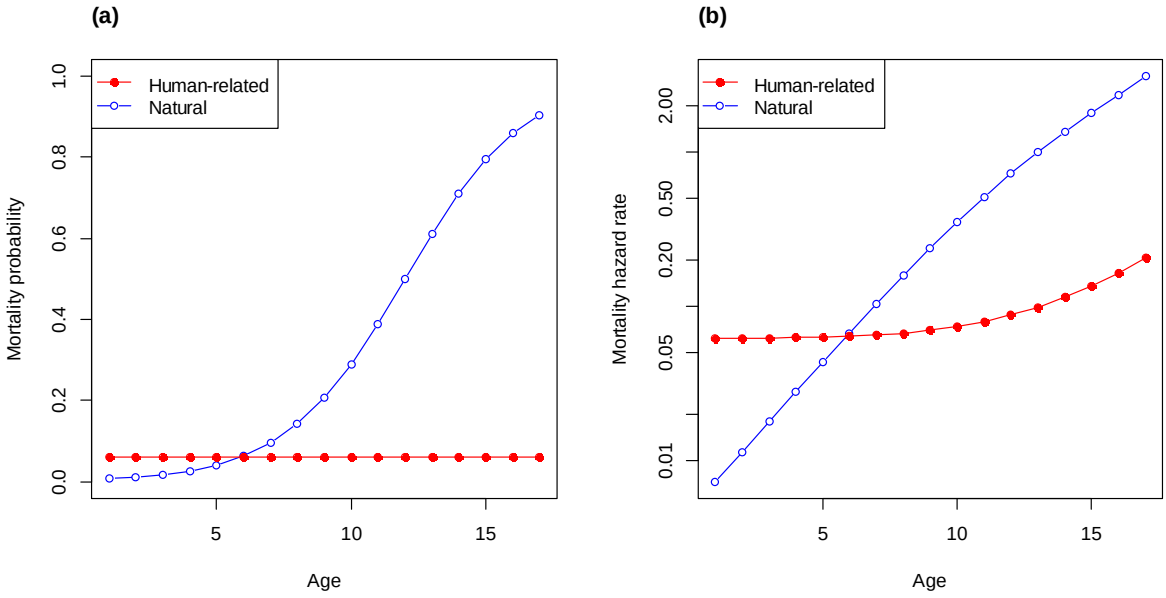


Figure 4. An example of two age related competing causes of mortality; human-related and non-human-related (natural). Mortality probabilities are plotted in panel a, while corresponding time-averaged mortality hazard rates are shown in panel b. The fact that one of the two mortality probabilities is constant (red line in panel a) while the other increases with age, implies that both cause specific mortality hazard rates must increase with age (panel b) – see equations in Table 1. Mortality hazard rates (panel b) are plotted on a logarithmic scale for easier comparison of relative differences. In this example, the mortality hazard rate for human related causes (red line in panel b) increases more than 3 times from age 1 to age 17 even though mortality probability remains constant. Compare figure with Figure 4 in Koons *et al.* (2014).

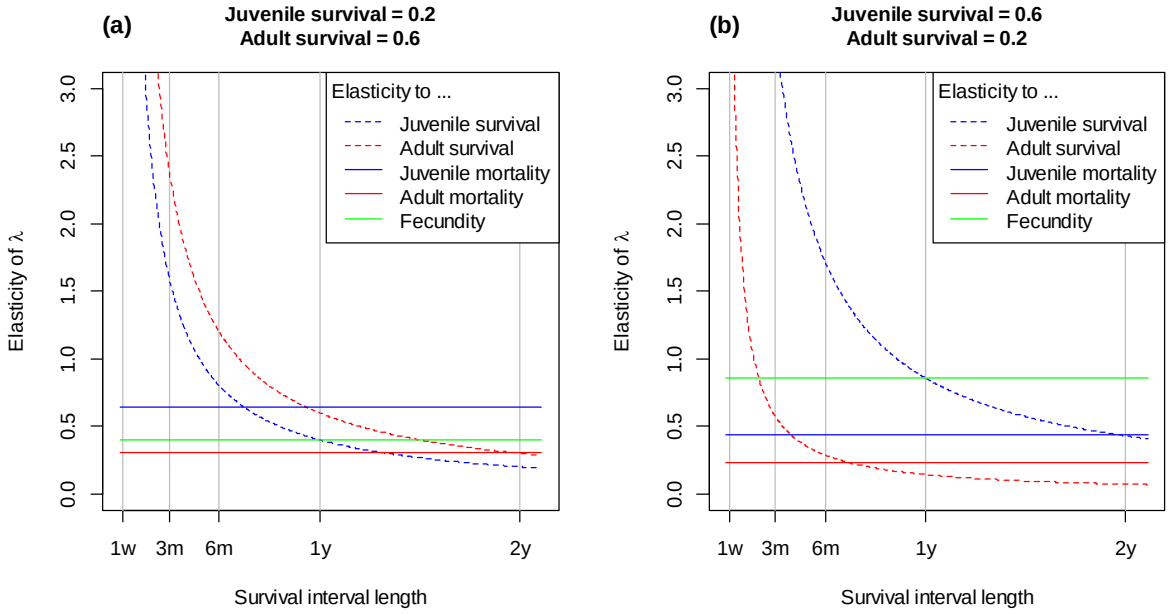


Figure 5. Elasticity of population growth rate (λ) to juvenile (blue lines) and adult (red lines) survival (dashed lines) and mortality hazard rate (solid lines), as well as to fecundity (green lines), when the model is parameterized with survival probabilities over different interval lengths (x-axis). The model assumes maturation at age 1 years and fecundity is 2 offspring per individuals for all ages. Vertical lines indicate 1 week, 3 months, 6 months, 1 year and 2 years. In panel (a), yearly juvenile survival is 0.2, while adult survival is 0.6. In panel (b), yearly juvenile survival is 0.6, while adult survival is 0.2. For mortality hazard rates, the negative elasticity (i.e., the elasticity of lambda to a *decrease* in mortality hazard rate, or an increase in "current life-expectancy" (Appendix S1)) is plotted for easier comparison.

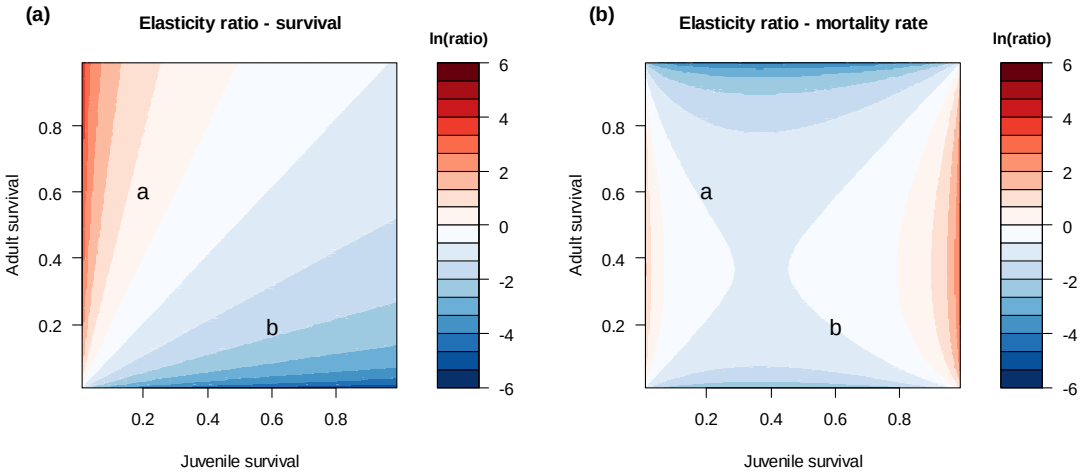


Figure 6. Elasticity of population growth rate (λ) to adult survival (panel a) and mortality hazard rate (panel *b*) relative to elasticity to respectively juvenile survival and juvenile mortality hazard rate. In both panels, the elasticity ratios (colour-key on a $\log_e$-scale) are shown for all possible combinations of juvenile (x-axis) and adult survival (y-axis). As in Figure 5, maturation occurs at age 1 and fecundity is 2 offspring per individuals for all ages. Plotted labels "a" and "b" correspond to the survival probabilities used in respectively panel (a) and panel (b) of Figure 5. It can be shown that the elasticity ratio with respect to mortality (panel b) is the hazard ratio $\left(\frac{\bar{h}_{adult}}{\bar{h}_{juvenile}}\right)$ times the elasticity ratio with respect to survival (panel a).

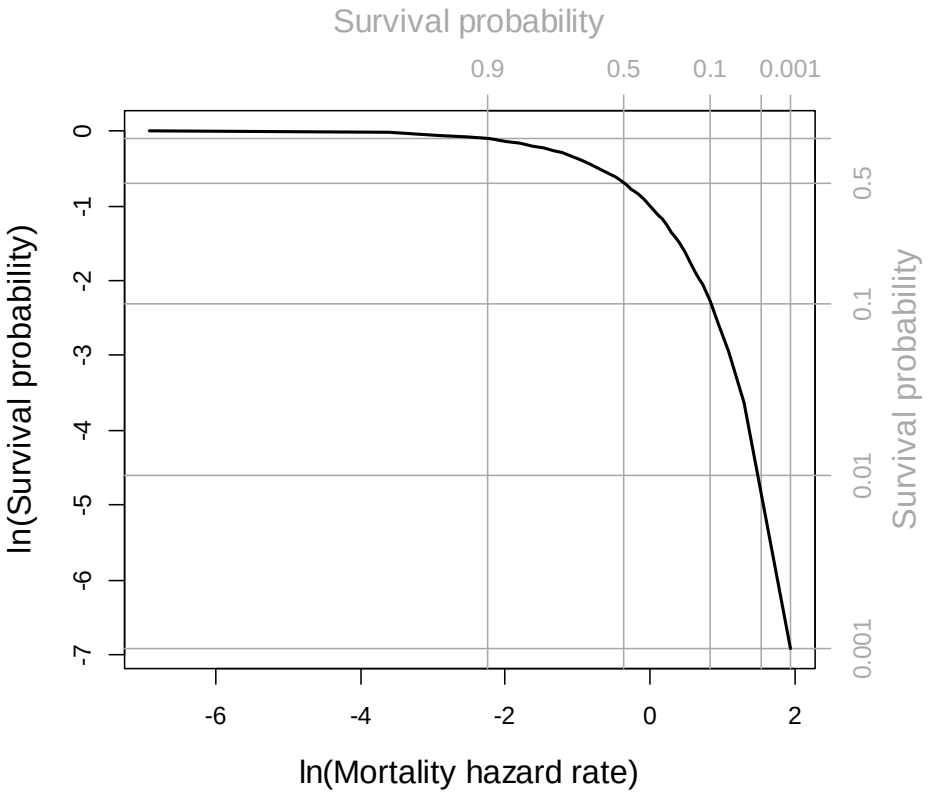


Figure 7. The natural logarithm of survival probability (y-axis) as a function of the natural logarithm of time-averaged mortality hazard rate (x-axis). The slope of the curve is the relative change in survival per relative change in mortality hazard rate. Corresponding survival probabilities are shown in the top and right axes and grey lines.

Table 1. Matrix of three alternative parameterizations of state transition probabilities in a multi-state model of two mutually exclusive and collectively exhaustive causes of mortality, assuming that cause specific hazard rates remain proportional within intervals of length 1. Definitions of the parameters are given in the diagonal cells. These parameters may be interval and individual specific, and may be modelled as random effects or as functions of covariates in a hierarchical model.

Parameterization	**State transition probabilities**[1]	**Survival and proportion died of cause 1**	**Mortality hazard rates**
State transition probabilities; $P_1 + P_2 + S = 1$ (Schaub & Lebreton 2004; Bischof *et al.* 2009; Servanty *et al.* 2010; Péron 2013; Koons *et al.* 2014)	$\boldsymbol{P_1}$ **= 'Probability of dying of cause 1'** $\boldsymbol{P_2}$ **= 'Probability of dying of cause 2'** $\boldsymbol{S = 1 - P_1 - P_2}$ **= 'Survival probability'**	$S = 1 - P_1 - P_2$ $\alpha = \frac{P_1}{P_1 + P_2}$	$\bar{h}_1 = -\log(1 - P_1 - P_2)\frac{P_1}{P_1 + P_2}$ $\bar{h}_2 = -\log(1 - P_1 - P_2)\frac{P_2}{P_1 + P_2}$
Survival and proportion died of cause 1; S and α (Schaub & Pradel 2004; Schaub 2009; Souchay & Schaub 2016)	$P_1 = (1 - S)\alpha$ $P_2 = (1 - S)(1 - \alpha)$ $S = S$	$\boldsymbol{S}$ **= 'Survival probability'** $\boldsymbol{\alpha}$ **= 'Proportion of total hazard rate attributable to cause 1', or 'Proportion died of cause 1'**[2]	$\bar{h}_1 = -\log(S)\,\alpha$ $\bar{h}_2 = -\log(S)\,(1 - \alpha)$
Mortality hazard rates; h_1 and h_2 (Appendix S2 - this paper)	$P_1 = \left(1 - e^{-(\bar{h}_1 + \bar{h}_2)}\right)\frac{\bar{h}_1}{\bar{h}_1 + \bar{h}_2}$ $P_1 = \left(1 - e^{-(\bar{h}_1 + \bar{h}_2)}\right)\frac{\bar{h}_2}{\bar{h}_1 + \bar{h}_2}$ $S = e^{-(\bar{h}_1 + \bar{h}_2)}$	$S = e^{-(\bar{h}_1 + \bar{h}_2)}$ $\alpha = \frac{\bar{h}_1}{\bar{h}_1 + \bar{h}_2}$	$\boldsymbol{\bar{h}_1}$ **= 'Time-averaged mortality hazard rate of cause 1'** $\boldsymbol{\bar{h}_2}$ **= 'Time-averaged mortality hazard rate of cause 2'**

[1] Parameters P_1, P_2 and S constrained to sum to 1 by e.g. the 'mlogit' link-function ($P_k = \frac{\exp(\eta_k)}{1 + \sum_{l=1}^{K} \exp(\eta_l)}$ and $S = 1 - \sum_{k=1}^{K} P_k$).

[2] In the three references α was interpreted as 'Proportion died of cause 1' which does not assume that hazard rates remain proportional within intervals.

Glossary

Mortality hazard rate: The latent intensity of deadly events that an individual is exposed to (Box 1). Synonymous to 'force of mortality' and 'failure rate'. Often referred to as just 'mortality rate'.

Survival probability: The probability that an individual survives an interval of a given length given that it is alive at the beginning of the interval. A unit-free measurement on an **absolute scale**. Often misleadingly referred to as 'survival rate'.

Mortality probability: The probability that an individual dies during a given time interval, equal to one minus the survival probability for the same interval. Often misleadingly referred to as 'mortality rate'.

Risk: Usually synonymous to **mortality probability**.

Odds: When applied to survival, **survival probability** divided by **mortality probability**.

Odds-ratio: One **odds** divided by another.

Hazard ratio: One **hazard rate** divided by another.

Logit-link: When applied to a probability, a link-function on the form log('odds') = 'linear predictor' $= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots = \boldsymbol{\beta x}$.

Loglog-link: When applied to survival probability, a link-function on the form log('time-averaged mortality hazard rate') = 'linear predictor' (see Box 1).

Permissible transformation: Transformations that preserve the relevant relationships among measurements (Houle *et al.* 2011)

Absolute-scaled variable: A scale type with no **permissible transformations** (e.g. probabilities).

Ratio-scaled variable: A positive variable with a true (nature-given) zero (zero refers to "non-existence" or "total absence"). Examples include age, mass and mortality hazard rates. The only **permissible transformations** are multiplications and divisions by strictly positive numbers.

Absolute difference/effect: A difference or effect quantified as one quantity minus another. E.g., the absolute difference between values 2 and 1 is 1.

Relative difference/effect: A difference or effect quantified as one quantity divided by another. E.g., the relative difference between values 2 and 1 is 2. The **absolute difference** on a logarithmic scale equals the logarithm of the relative difference since $\log(a) - \log(b) = \log(a/b)$.

Linear contrast: The **absolute difference** between two predicted values at the scale of the link-function.

Log-linear models: Models on the form log('prediction') $= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots$. When the x's are log-transformed values of a predictor variable, a **relative difference** in the predictor variable leads to a **relative difference** in the predictions (e.g., a 10% increase in the predictor variable x_1 leads to a relative difference in the prediction equal to 1.1^{β_1}).

References

Benton, T. & Grant, A. (1999) Elasticity analysis as an important tool in evolutionary and population ecology. *Trends Ecol Evol.,* **14,** 467-471.

Bischof, R., Swenson, J.E., Yoccoz, N.G., Mysterud, A. & Gimenez, O. (2009) The magnitude and selectivity of natural and multiple anthropogenic mortality causes in hunted brown bears. *J Anim Ecol,* **78,** 656-665.

Boyce, M.S., Sinclair, A.R.E. & White, G.C. (1999) Seasonal compensation of predation and harvesting. *Oikos,* **87,** 419-426.

Caswell, H. (2001) *Matrix Population Models; Construction, analysis and interpretation.* Sinauer Associates, Sunderland, Massachusetts.

Caswell, H. (2007a) Extrinsic mortality and the evolution of senescence. *Trends Ecol Evol,* **22,** 173-174.

Caswell, H. (2007b) Sensitivity analysis of transient population dynamics. *Ecol Lett,* **10,** 1-15.

Charlesworth, B. (1980) *Evolution of Age-Structured Populations*. Cambridge University Press.

Choquet, R., Garnier, A., Awuve, E. & Besnard, A. (2017) Transient state estimation using continuous-time processes applied to opportunistic capture–recapture data. *Ecological Modelling,* **361,** 157-163.

Choquet, R., Rouan, L. & Pradel, R. (2009) Program E-SURGE: a software application for fitting multievent models. *Modeling Demographic Processes in Marked Populations* (eds D.L. Thompson, E.G. Cooch & M.J. Conroy). Springer.

Collett, D. (2014) *Modelling survival data in medical research,* Third edn. CRC Press.

Conn, P.B., Cooch, E.G. & Caley, P. (2012) Accounting for detection probability when estimating force-of-infection from animal encounter data. *Journal of Ornithology,* **152,** S511-S520.

de Valpine, P., Scranton, K., Knape, J., Ram, K. & Mills, N.J. (2014) The importance of individual developmental variation in stage-structured population models. *Ecol Lett,* **17,** 1026-1038.

Ergon, T., Yoccoz, N. & Nichols, J.D. (2009) Estimating latent time of maturation and survival costs of reproduction in continuous time from capture-recapture data. *Modeling Demographic Processes in Marked Populations* (eds D. Thomson, E.G. Cooch & M.J. Conroy). Springer Science+Business Media, LLC.

Gaillard, J.M., Loison, A., Villefont, A. & Festa-Bianchet, M. (2004) Assesing senescence patterns in populations of large mammals. *Animal Biodiversity and Conservation,* **27,** 47-58.

Gaillard, J.M. & Yoccoz, N.G. (2003) Temporal variation in survival of mammals: A case of environmental canalization? *Ecology,* **84,** 3294-3306.

Houle, D., Pélabon, C., Wagner, G.P. & Hansen, T.F. (2011) Measurement and meaning in biology. *The quarterly review of biology,* **86,** 3-34.

Keyfitz, N. & Caswell, H. (2005) *Applied Mathematical Demography*. Springer.

Koons, D.N., Gamelon, M., Gaillard, J.-M., Aubry, L.M., Rockwell, R.F., Klein, F., Choquet, R. & Gimenez, O. (2014) Methods for studying cause-specific senescence in the wild. *Methods in Ecology and Evolution,* **5,** 924-933.

Laake, J.L., Johnson, D.S. & Conn, P.B. (2013) marked: An R package for maximum-likelihood and MCMC analysis of capture-recapture data. *Methods in Ecology and Evolution,* **4,** 885-890.

Lebreton, J.-D., Burnham, K.P., Clobert, J. & Anderson, D.R. (1992) Modelling survival and testing biological hypotheses using marked animals: a unified approach with case studies. *Ecological Monographs,* **62,** 67-118.

Link, W.A. & Doherty, P.F. (2002) Scaling in sensitivity analysis. *Ecology,* **83,** 3299-3305.

Lunn, M. & McNeil, D. (1995) Applying Cox regression to competing risks. *Biometrics,* **51,** 524–532.

McCrea, R.S. & Morgan, B. (2015) *Analysis of capture-recapture data*. Chapman & Hall/CRC.

McDonald, J.L., Franco, M., Townley, S., Ezard, T.H.G., Jelbert, K. & Hodgson, D.J. (2017) Divergent demographic strategies of plants in variable environments. *Nat Ecol Evol,* **1,** 29.

Miller, T.J. & Andersen, P.K. (2008) A finite-state continuous-time approach for inferring regional migration and mortality rates from archival tagging and conventional tag-recovery experiments. *Biometrics,* **64,** 1196-1206.

Nichols, J.D., Conroy, M.J., Anderson, D.R. & Burnham, K.P. (1984) Compensatory mortality in waterfowl populations: a review of the evidence and implications for research and

management. *Transactions of North American Wildlife and Natural Resources Conference,* **49,** 535-554.
Péron, G. (2013) Compensation and additivity of anthropogenic mortality: life-history effects and review of methods. *J Anim Ecol,* **82,** 408-417.
Quinn, T.J. & Deriso, R.B. (1999) *Quantitative fish dynamics*. Oxford University Press.
Royle, J.A. & Dorazio, R.M. (2008) *Hierichical modelling and inference in ecology*. Academic Press.
Schaub, M. (2009) Evaluation of bias, precision and accuracy of mortality cause proportion estimators from ring recovery data. *Modeling Demographic Processes in Marked Populations* (eds D. Thomson, E.G. Cooch & M.J. Conroy). Springer Science+Business Media, LLC.
Schaub, M. & Lebreton, J.D. (2004) Testing the additive versus the compensatory hypothesis of mortality from ring recovery data using a random effects model. *Animal Biodiversity and Conservation,* **27,** 73-85.
Schaub, M. & Pradel, R. (2004) Assessing the relative importance of different sources of mortality from recoveries of marked animals. *Ecology,* **85,** 930.
Servanty, S., Choquet, R., Baubet, E., Brandt, S., Gaillard, J.-M., Schaub, M., Toigo, C., Lebreton, J.-D., Buoro, M. & Gimenez, O. (2010) Assessing whether mortality is additive using marked animals: a Bayesian state–space modeling approach. *Ecology,* **91,** 1916-1923.
Souchay, G. & Schaub, M. (2016) Investigating rates of hunting and survival in declining european lapwing populations. *PLoS One,* **11,** e0163850.
van Tienderen, P. (2000) Elasticities and the link between demographic and evolutionary dynamics. *Ecology,* **81,** 666-679.
White, G.C. & Burnham, K.P. (1999) Program MARK: Survival estimation from populations of marked animals. *Bird Study,* **46 Supplement,** 120-138.
Wolfe, M.L., Koons, D.N., Stoner, D.C., Terletzky, P., Gese, E.M., Choate, D.M. & Aubry, L.M. (2015) Is anthropogenic cougar mortality compensated by changes in natural mortality in Utah? Insight from long-term studies. *Biological Conservation,* **182,** 187-196.

Supporting information for 'Study mortality with hazard rates, not probabilities'

Torbjørn Ergon[a], Ørnulf Borgan[b], Chloe Rebecca Nater[a] and Yngvild Vindenes[a]*

[a] *Centre for Ecological and Evolutionary Synthesis, University of Oslo, P.O. Box 1066 Blindern, 0316 Oslo, Norway.*

[b] *Department of Mathematics, University of Oslo, P.O. Box 1066 Blindern, 0316 Oslo, Norway.*

[*] *Corresponding author, e-mail t.h.ergon@ibv.uio.no*

"2018-04-08"

Abstract

This document provides additional information relating to each section and figure in the main article (https://doi.org/10.1101/216739 (link will be updated upon acceptance in the journal)). The document is written in R Markdown. The Markdown source file (.Rmd) is available at https://github.com/torbjore/study_mortality_with_hazard_rates. It contains all R-code needed to reproduce all figures and examples in this document (see https://rmarkdown.rstudio.com/).

Contents

Appendix S1: Details on sections and figures **3**
1.1 Introduction (Box 1) . . . 3
1.1.1 The relationship between mortality hazard rates and survival probabilities . . . 3
1.2 Logit-linear models of survival vs. log-linear models of mortality rates . . . 3
1.2.1 Relationships . . . 3
1.2.2 Figure 1 . . . 5
1.2.2.1 Code - Figure 1 . . . 5
1.2.3 Figure 2 . . . 7
1.2.3.1 Code - Figure 2 . . . 7
1.3 Multiple sources of mortality and competing risks . . . 8
1.3.1 Figure 3 . . . 9
1.3.1.1 Code - Figure 3 . . . 9
1.3.2 Figure 4 . . . 10
1.3.2.1 Code - Figure 4 . . . 10
1.4 Elasticity analyses based on mortality hazard rates instead of survival probabilities . . . 12
1.4.1 Relationships . . . 12
1.4.1.1 Elasticity to S and $\bar{h}$ . . . 12
1.4.1.2 Dependencies on interval length . . . 12
1.4.1.3 Elasticity to mortality probability . . . 12
1.4.1.4 Elasticity to "current life-expectancy" . . . 12
1.4.1.5 Perspectives on hazard rates and models of stage-duration distributions . . . 13
1.4.2 Figures 5 and 6 . . . 14
1.4.2.1 Code - Figure 5 . . . 14
1.4.2.2 Code - Figure 6 . . . 17
1.4.3 Figure 7 . . . 19
1.4.3.1 Code - Figure 7 . . . 19
1.4.4 Variance-stabilized sensitivities . . . 20

Appendix S2: Competing risks and cause-specific mortality probabilities **22**

Appendix S3: Fitting capture-recapture models with cause-specific mortality **24**
3.1 Introduction 24
3.2 Preparation 25
3.3 Example 25

References **33**

Appendix S1: Details on sections and figures

The headings below are the same as in the article. Each section contains explanations and R-code for the figures referred to under each heading.

1.1 Introduction (Box 1)

1.1.1 The relationship between mortality hazard rates and survival probabilities

In Box 1 we explained the relationship between individual mortality hazard rates and survival probabilities (eq. (1) in the main text) by considering an individual point process for deadly events. We chose this approach to give an intuitive understanding of the hazard rate as an "intensity of deadly events" that individuals are exposed to and to set the stage for explaining the issue of competing risks in an intuitive manner in Appendix S2. The following derivation follows the derivation in Quinn & Deriso (1999, p. 10), referring to Ricker (1975), only that we here treat the mortality hazard rate as an individual trait and not a population level characteristic.

Consider a hypothetical population of $N(t_1)$ identical individuals having identical expectations for their future environment at time t_1. If there is no reproduction, the instantaneous rate of change in number of individuals, $N(t)$, will be

$$\frac{\partial N(t)}{\partial t} = -hN(t)$$

where h is the mortality hazard rate in units of t^{-1}, here considered to be constant. Letting Δ represent a certain increment in time and solving for $N(t_1 + \Delta)$ as a function of $N(t_1)$ gives

$$N(t_1 + \Delta) = N(t_1)e^{-h\Delta}$$

To convert this to a survival probability over the interval, we divide by $N(t_1)$,

$$S_{t_1 \to t_1 + \Delta} = \frac{N(t_1 + \Delta)}{N(t_1)} = e^{-h\Delta}$$

If the mortality hazard rate is considered as a continuous function of time, $h(t)$, we can calculate the survival probability as a product of many infinitesimally small sections of the interval from time t_1 to t_2 to obtain eq. (1) in the main text,

$$S_{t_1 \to t_2} = e^{-\int_{t_1}^{t_2} h(t)dt}.$$

For further relationships and derivations, see Collet (2014, chap. 1.3) and section 1.4.1.4 below.

1.2 Logit-linear models of survival vs. log-linear models of mortality rates

1.2.1 Relationships

Survival probabilities, S, in any discrete-time model can be replaced by $S = \exp(-\bar{h})$, where $\bar{h}$ is the time-averaged mortality hazard rate in the given time interval(s). The unit of measurement for $\bar{h}$ is here the inverse of the length of the interval (see Box 1 in the main text). That is, if S is the probability of surviving one year, the unit of $\bar{h}$ is year^{-1}, whereas if S is the probability of surviving one month, the unit of $\bar{h}$ is month^{-1}.

Since $\bar{h}$ is a ratio-scaled intensity, it is natural to model multiplicative (proportional) effects on $\bar{h}$ through a log-link,

$$\bar{h} = \exp(\eta),$$

where η is a linear predictor of explanatory variables,

$$\eta = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ...$$

This corresponds to using a loglog link-function on survival probabilities,

$$S = \exp(-\bar{h}) = \exp(-\exp(\eta))$$

The exponent of contrasts in the linear predictor (e.g. representing different age groups or individuals with different body mass or sex) then becomes hazard ratios[i],

$$\exp(\eta_2 - \eta_1) = \frac{\exp(\eta_2)}{\exp(\eta_1)} = \frac{\bar{h}_2}{\bar{h}_1}.$$

One could reparameterize the discrete-time model with e.g. monthly survival probabilities (S_m) instead of yearly survival probabilities (S_y). If one assumes that all months of the year have the same survival probability, this amounts to just replacing all S_y in the model with $(S_m)^{12}$.[ii] When using a loglog-link on survival probabilities, the model is invariant to such replacements because the time-averaged hazard rate remains unchanged. If this does not appear as obvious, one may see this better by first calculating the time-averaged hazard rate as $\bar{h}_y = -\log(S_y)$. The unit for $\bar{h}_y$ is here year^{-1}. Similarly, the time-averaged hazard rate can also be calculated as $\bar{h}_m = -\log(S_m)$, where $\bar{h}_m$ is expressed in units of month^{-1}. Since $S_y = (S_m)^{12}$, we see that $\bar{h}_y = -\log((S_m)^{12}) = -12\log(S_m) = 12\bar{h}_m$, and since an intensity of one event per month is the same as an intensity of 12 events per year, the two hazard rates $\bar{h}_y$ and $\bar{h}_m$ are identical (they just have different units of expression). Note also that the ratio of two hazard rates (i.e., hazard ratios) are invariant to which unit of expression is used for the hazard rates. Hence, when modelling survival probabilities through a loglog-link, all parameters in the linear predictor except the intercept are invariant to reparameterizations on the form $S_y = (S_m)^{12}$. This can be seen from expanding $\bar{h}_y = 12\bar{h}_m$,

$$\bar{h}_y = 12\bar{h}_m = 12\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + ...) = \exp((\log(12) + \beta_0) + \beta_1 x_1 + \beta_2 x_2 + ...)$$

It is (unfortunately) common practice to model effects of covariates on survival probabilities through the use of a logit link-function. That is, survival probability is modelled as

$$S = \frac{\exp(\eta)}{1 + \exp(\eta)},$$

where η is the linear predictor. The exponent of the linear predictor then becomes the survival-odds,

$$\exp(\eta) = \frac{S}{1 - S}.$$

Just like survival probabilities, but unlike hazard rates, survival-odds are unit-free absolute scale measurements that cannot be transformed in any way without losing their original meaning.

If we fit an effect of a time-invariant covariate (e.g., an individual covariate) on monthly survival using a logit-link, the relationship between the covariate and yearly survival will be

$$S_y = \left(\frac{\exp(\eta)}{1 + \exp(\eta)}\right)^{12}$$

[i] Note that some computer programs for analyses of capture-recapture data, such as MARK, define the inverse loglog-link as $S = \exp(-\exp(-\eta))$, in which case the hazard ratio $\frac{\bar{h}_2}{\bar{h}_1}$ becomes $\exp(\eta_1 - \eta_2)$.

[ii] Note that this is a reparameterization since S_y and S_m have different meanings - it is not about "selecting a time-scale for the survival probabilities". Survival probabilities are unit-free absolute scale measurements that cannot be "re-scaled" in any way. Such reparameterizations of the general (full, unconstrained) model may be specified in e.g. MARK, and is commonly done when time between capture occasions vary over the study period.

which is not easily simplified. Hence, replacing S_y with $(S_m)^{12}$ changes the functional form of the model when survival probabilities are modelled by a logit-link involving continuous covariates. In contrast, when using a loglog-link, we obtain

$$S_y = (\exp(-\bar{h}_m))^{12} = (\exp(-\exp(\eta)))^{12} = \exp(-12\exp(\eta)) = \exp(-\bar{h}_y),$$

and, as also explained above, the functional relationships between the covariates and survival remains unchanged (only the intercept is changed to reflect the change in unit of expression for the time-averaged hazard rates).

In conclusion, when using the model $S = (\text{logit}^{-1}(\beta_0 + \beta_1 x_1))^n$ where n is the length of the survival interval in the chosen time-units, n can be interpreted as a parameter that is fixed by the modeller (e.g., by setting the time-interval length in MARK). In contrast, the model $S = (\text{loglog}^{-1}(\beta_0 + \beta_1 x_1))^n$ is equivalent to $S = (\text{loglog}^{-1}(\beta_0' + \beta_1 x_1))$ where $\beta_0' = \log(n) + \beta_0$. Hence, when using a loglog-link, changing the time-scale just amounts to changing the definition of the intercept parameter but the model remains the same (i.e., it is a reparameterization).

1.2.2 Figure 1

Figure 1 in the main text illustrates the effect of replacing S_y with $(S_n)^n$ where $1/n$ is the length of the time-intervals used in the discrete model. To facilitate comparison of functional shapes, the intercepts and slopes of logit survival for all curves are defined such that yearly survival probability equals 0.05 when the covariate value is -2 and 0.95 when the covariate value is 2. If the models were fitted to data, the functional forms would be the same as in the figure, and predicted values would be in the same range, but the curves would not necessarily cross at the same place.

1.2.2.1 Code - Figure 1

```
# Defining functions
logit = function(S) log(S/(1-S))
invlogit = function(eta) exp(eta)/(1+exp(eta))
loglog = function(S) log(-log(S))
invloglog = function(eta) exp(-exp(eta))
rescaled.invlogit = function(x, il, s1=0.05, s2 = 0.95){ # il = 'interval length'
  a = mean(c(logit(s2^il), logit(s1^il))) # intercept
  b = (logit(s2^il)-logit(s1^il))/4       # slope
  Sy = invlogit(a + b*x)^(1/il)           # yearly survival
  return(Sy)
}
rescaled.invloglog = function(x, il, s1=0.05, s2 = 0.95){ # il = 'interval length'
  a = mean(c(loglog(s2^il), loglog(s1^il))) # intercept
  b = (loglog(s2^il)-loglog(s1^il))/4       # slope
  Sy = invloglog(a + b*x)^(1/il)            # yearly survival
  return(Sy)
}

# Values for plotting
x = seq(-3,3,length.out=60)
Syy.logit = rescaled.invlogit(x, 1)      # Using logit-link applied to yearly survival
Sym.logit = rescaled.invlogit(x, 1/12)   # Using logit-link applied to monthly survival
Sy12.logit = rescaled.invlogit(x, 12)    # Using logit-link applied to 12-year survival
Syy.loglog = rescaled.invloglog(x, 1)    # Using loglog-link applied to yearly survival
Sym.loglog = rescaled.invloglog(x, 1/12) # Using loglog-link applied to monthly survival
```

```r
# All Syy.loglog == Sym.loglog (but round-off errors occur, hence using round())
all(round(Syy.loglog, 12) == round(Sym.loglog, 12))
```

```
## [1] TRUE
```

```r
# Making plot
plot(x, Syy.logit, type="l", col="blue",
     xlab="Time-invariant covariate", ylab="Yearly survival")
lines(x, Sym.logit, col="red")
lines(x, Sy12.logit, col="green")
lines(x, Syy.loglog, lty=2)
abline(v=-2, lty=3)
abline(v=2, lty=3)
legend("topleft",
  legend = c(
    expression(logit(italic(S)[monthly])),
    expression(logit(italic(S)[yearly])),
    expression(logit(italic(S)[12-yearly])),
    expression(loglog(italic(S^{n})))
  ),
  lty=c(1,1,1,2), col=c("red", "blue", "green", "black"), bg="white"
)
```

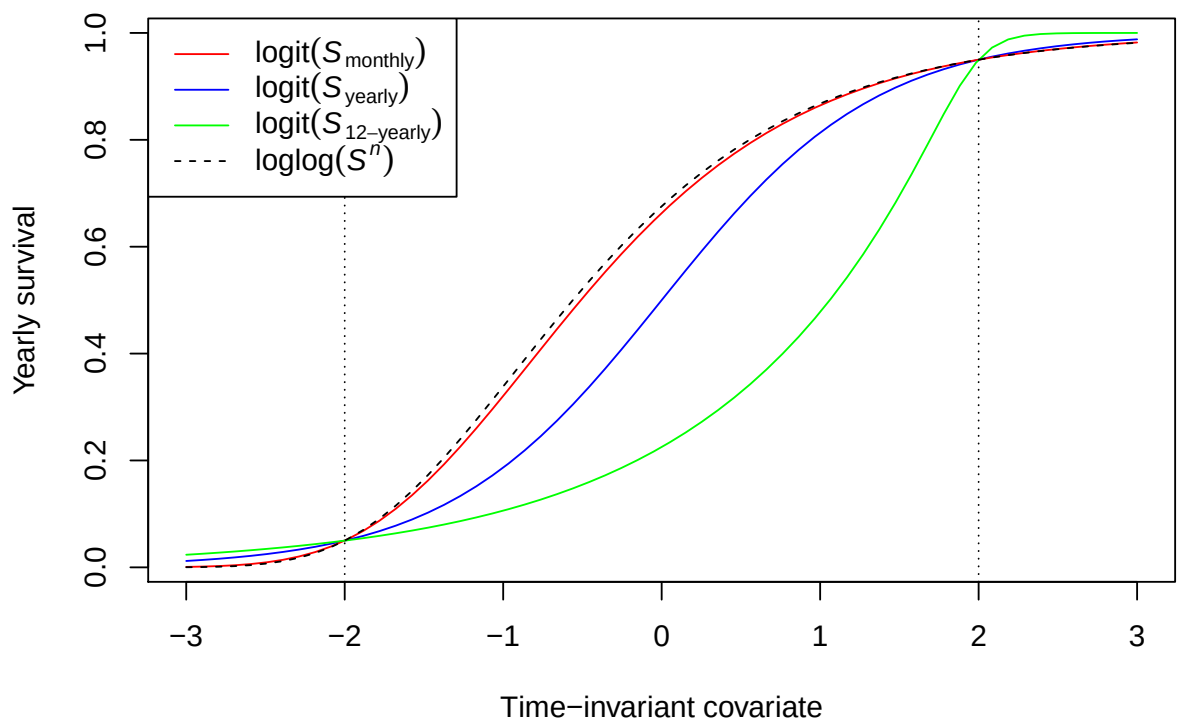


Figure 1: See caption in the main text.

1.2.3 Figure 2

Odds-ratios are hard to interpret, not only because their meaning depend on the chosen interval length of the discrete-time model, but also because, for a given hazard ratio, they depend on the survival probability of the reference group. This is illustrated in Figure 2 in the main text.

1.2.3.1 Code - Figure 2

```
# Function for computing odds-ratio for a given hazard ratio (hr) and survival
# probability of reference group (s1)
OR = function(s1, hr){
  s2 = exp(-(-log(s1)*hr))
  (s2/(1-s2))/(s1/(1-s1))
}
ORv = Vectorize(OR)

# Constants
scale = seq(0, 2.1, length.out=60)      # Survival interval length in years (x-axis)
s = c(0.01, 0.1, 0.25, 0.5, 0.75, 0.99) # Survival probabilities of reference
HR = 0.5                                # Hazard ratio to be used in plot

# Draw the plot
Col = rainbow(length(s))
Col[2] = "#FF8000FF" # Changing the light yellow to orange
plot(1,1, type="n", ylim=c(0,6), xlim=range(scale), axes=F,
     xlab="Survival interval length", ylab="Odds-ratio")
box()
axis(2)
At = c(7/365, 0.5, 1, 2)
Lab = c("1w", "6m", "1y", "2y")
axis(1, at = At, labels=Lab)
abline(v = At, col="grey")
abline(h=1, lty=2)
lines(range(scale), c(HR,HR), lwd=2) # Line for the hazard ratio
for(i in 1:length(s)){
  lines(scale, ORv(s[i]^scale, HR), col=Col[i], lwd=1.5)
}
legend("topleft", legend=s, col=Col, lty=1, bg="white", lwd=1.5,
       title="   Yearly survival of reference   ", ncol=2)

# Adding points corresponding to the examples in the text
points(c(1,2), ORv(c(0.25, 0.25^2), HR), cex=1.7)     # Original environment; odds-ratio
points(c(1,2), c(HR,HR), cex=1.7)                     # Original environment; hazard-ratio
points(c(1,2), ORv(c(0.5, 0.5^2), HR), pch=4, cex=1.7)# Improved environment; odds-ratio
points(c(1,2), c(HR,HR), pch=4, cex=1.7)              # Improved environment; hazard ratio
```

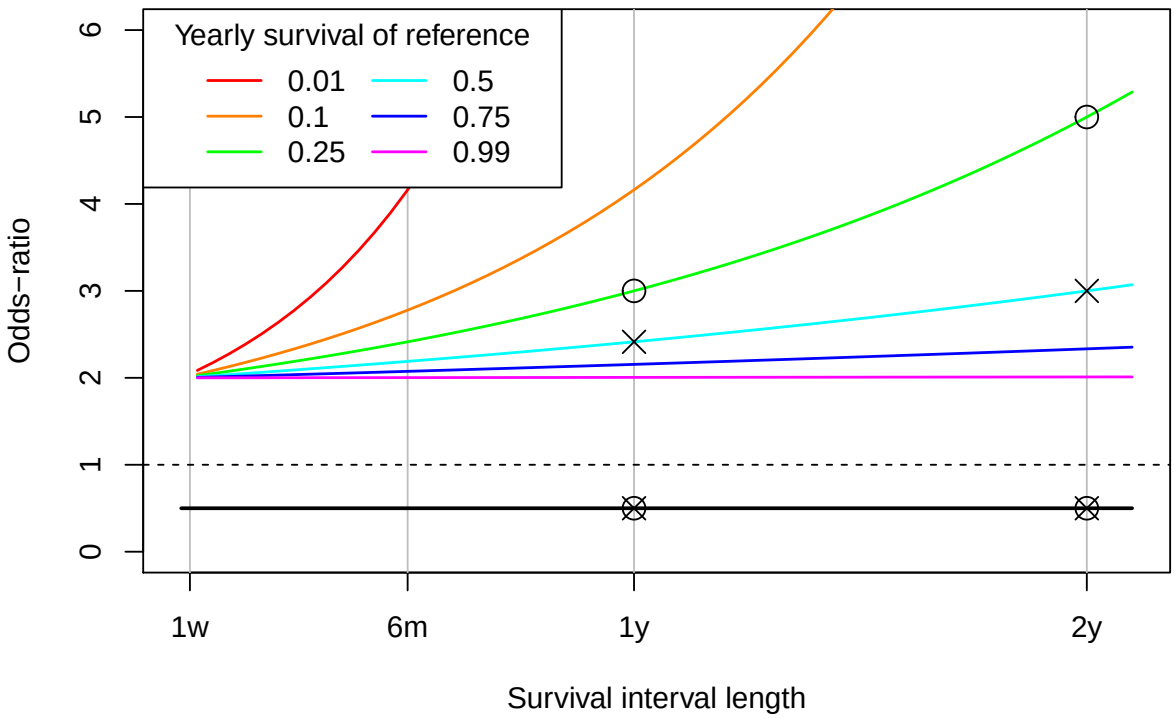


Figure 2: See caption in the main text.

From Fig. 2, it may be noted that odds-ratios approach the inverse of the hazard ratio ($1/0.5 = 2$) when survival probabilities are high; i.e., when the survival probability of the reference is high (magenta line) or when interval lengths are short (to the left in the plot).

1.3 Multiple sources of mortality and competing risks

Mortality hazard rates are more elucidating with respect to the underlying mechanisms than mortality probabilities when studying associations between two or more sources of mortality in a population because mortality probabilities are intrinsically dependent; increasing the hazard rate for one cause of mortality will automatically reduce the probabilities of dying from all other causes (Appendix S2). Figure 3 in the main text shows how this negative intrinsic dependency can dominate empirical correlations between cause specific mortality probabilities across years even when correlations between the cause specific mortality hazard rates are strongly positive.

Figure 4 illustrates the relationships between cause-specific mortality hazard rates and survival probabilities in a age-dependent model. We note here that, in empirical studies of senescence, one should also consider potential biases arising from un-modelled individual heterogeneity (i.e., "frailty"[iii]) (Vaupel, Manten

[iii] Unobserved individual heterogeneity in mortality hazard rates. In the presence of frailty, the population will be increasingly dominated by individuals with low mortality hazard rates at higher age-classes. Hence, age-related patterns in population level mortality rates are not representative of individual level patterns.

& Stallard 1979; Cam et al. 2002; Collett 2014), temporary emigration (Langtimm 2009) and age-dependent dispersal (Ergon & Gardner 2014).

1.3.1 Figure 3

Figure 3 is produced by simulation. First, 10^5 values of log hazard rates, $\log(h_1)$ and $\log(h_2)$, were drawn from a bi-variate normal distribution with means given by the the colour legend ($\text{median}(h) = \exp(\text{mean}(\log(h)))$), standard deviations given above each panel, and correlations given by the x-axis. Cause-specific mortality probabilities where then computed from each of the 10^5 samples, and the Pearson correlation coefficients between them where then computed and plotted on the y-axis.

1.3.1.1 Code - Figure 3

```
library(MASS) # for 'mvrnorm'

# Functions
sim.cor = function(sigma, mu){
    Rho = seq(-1,1,length.out=50)
    cor.h = cor.P = median.S = rep(NA, length(Rho))
    for(i in 1:length(Rho)){
        rho = Rho[i] # correlation between log hazards
        S = matrix(c(
          sigma[1]^2, rho*sigma[1]*sigma[2],
          rho*sigma[1]*sigma[2], sigma[2]^2), 2,2, byrow=T)
        log.h = mvrnorm(100000, c(mu,mu), S)
        h = exp(log.h)
        # Cause-specific mortality probabilities from Table 1 in main text:
        P1 = (1-exp(-(h[,1]+h[,2])))*h[,1]/(h[,1]+h[,2])
        P2 = (1-exp(-(h[,1]+h[,2])))*h[,2]/(h[,1]+h[,2])
        cor.P[i] = cor(P1,P2)
    }
    return(list(Rho=Rho, cor.P=cor.P))
}

plot.panel = function(sigma){
  H = c(0.01, 0.1, 0.5, 1, 2)
  Mu = log(H)
  Col = rainbow(length(Mu))
  plot(0,0, type="n", xlab=expression(cor(log~italic(h[1]),~log~italic(h[2]))),
       ylab=expression(cor(italic(P[1]),~italic(P[2]))))
  abline(0,1, col="grey")
  abline(v=c(-1,0,1), col="grey")
  abline(h=c(-1,0,1), col="grey")
  for(i in 1:length(Mu)){
    X = sim.cor(rep(sigma,2), mu=Mu[i])
    lines(X$Rho, X$cor.P, col=Col[i])
  }
  legend("topleft", lty=1, col=Col, legend=H, title=expression(Median~italic(h)),
         bg="white")
}

# Make each panel
par(mfrow=c(1,3))
```

```
plot.panel(1)
title(expression(SD(log~italic(h)) == 1))
plot.panel(.1)
title(expression(SD(log~italic(h)) == 0.1))
plot.panel(.01)
title(expression(SD(log~italic(h)) == 0.01))
```

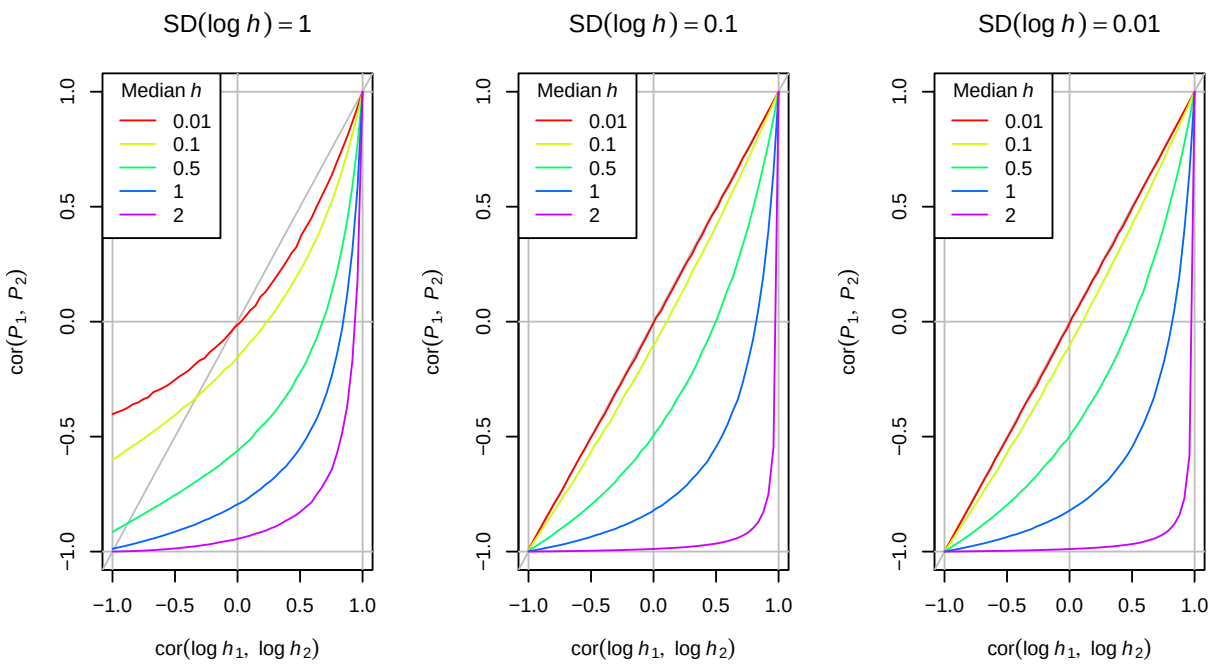


Figure 3: See caption in the main text.

1.3.2 Figure 4

Figure 4 illustrates the fact that, if one of two cause specific mortality *probabilities* is constant (red line in panel a) while the other increases with age, both cause specific mortality *hazard rates* must increase with age (panel b).

1.3.2.1 Code - Figure 4

```
# Parameters chosen to resemble Fig. 4 in Koons et al. 2014
x = 1:17 # Age (x-axis values)
c = 12 # Age at 50% survival
b = .45 # Slope

# Cause-specific mortality probabilities
P1 = rep(0.06, length(x)) # Human related
P2 = 1/(1+exp(-(b*(x-c)))) # Natural

# Corresponding time-averaged mortality hazard rates from Table 1 in main text
m1 = -log(1-P1-P2)*P1/(P1+P2)
```

```r
m2 = -log(1-P1-P2)*P2/(P1+P2)

# Plotting the two panels:
par(mfrow=c(1,2))

plot(mean(x),mean(P2), xlim=range(x), ylim=c(0,1), type="n",
     ylab="Mortality probability", xlab="Age")
lines(x,P2, col="blue")
points(x,P2, col="blue", pch=21, bg='white')
lines(x,P1, col="red")
points(x,P1, col="red", pch=19)
legend("topleft", legend=c("Human-related", "Natural"), col=c("red","blue"),
       pch=c(19,21), pt.bg = 'white', lty=1)
title("(a)", adj=0)

plot(mean(x),mean(m2), xlim=range(x), ylim=range(c(m1,m2)), type="n", log="y",
     ylab="Mortality hazard rate", xlab="Age")
lines(x,m2, col="blue")
points(x,m2, col="blue", pch=21, bg='white')
lines(x,m1, col="red")
points(x,m1, col="red", pch=19)
legend("topleft", legend=c("Human-related", "Natural"), col=c("red","blue"),
       pch=c(19,21), pt.bg='white', lty=1)
title("(b)", adj=0)
```

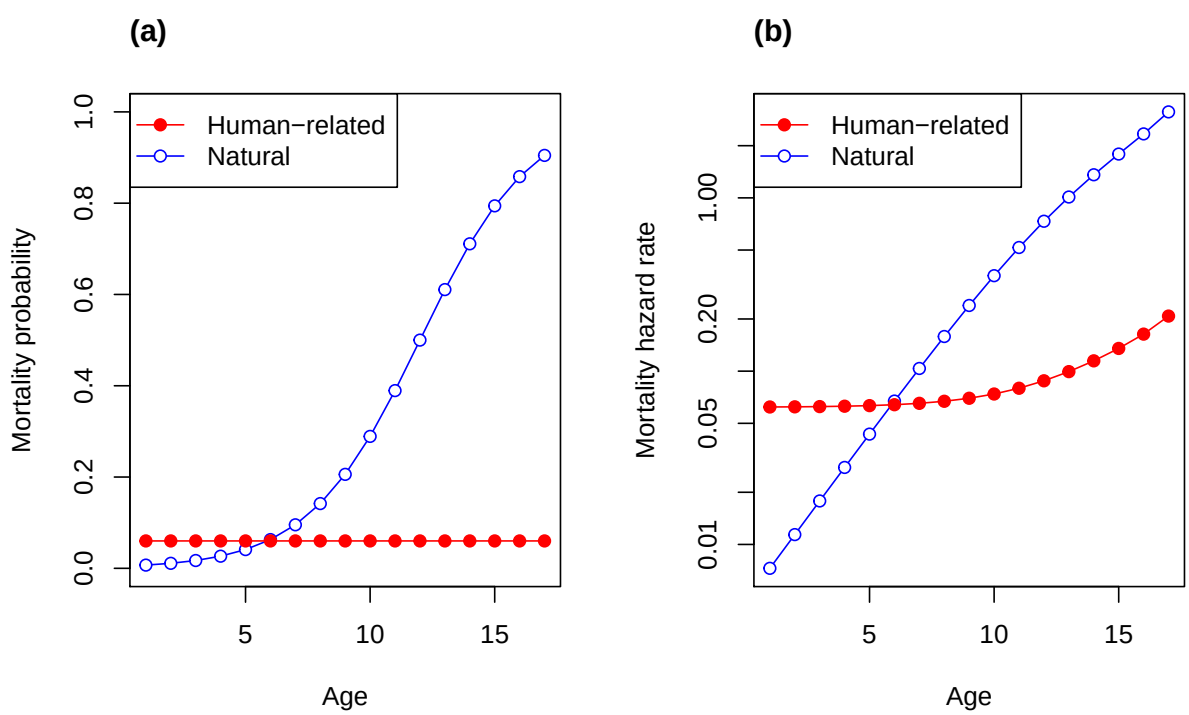


Figure 4: See caption in the main text.

1.4 Elasticity analyses based on mortality hazard rates instead of survival probabilities

1.4.1 Relationships

1.4.1.1 Elasticity to S and $\bar{h}$

In the main text, the relationship between elasticity of population growth rate λ to survival probability S over an interval, ϵ_S, and elasticity to time-averaged mortality hazard rate $\bar{h}$ during the same interval, $\epsilon_{\bar{h}}$, is given as

$$\epsilon_{\bar{h}} = -\bar{h}\epsilon_S$$

To derive this, note the relationship $\bar{h} = -\log(S)$ (eq. (2) in the main text), and remember that the elasticities are defined as $\epsilon_{\bar{h}} = \frac{\partial \log(\lambda)}{\partial \log(\bar{h})}$ and $\epsilon_S = \frac{\partial \log(\lambda)}{\partial \log(S)}$. By using the chain rule of derivation, we get

$$\epsilon_{\bar{h}} = \frac{\partial \log(\lambda)}{\partial \log(\bar{h})} = \frac{\partial \log(\lambda)}{\partial \log(S)} \frac{\partial \log(S)}{\partial \log(\bar{h})} = \epsilon_S \frac{\partial(-\bar{h})}{\partial \log(\bar{h})} = \epsilon_S(-\bar{h}) = \epsilon_S \log(S).$$

1.4.1.2 Dependencies on interval length

If the above survival probability applies to one year, S_y, one may choose to reparameterize the model with monthly survival probabilities, S_m, assuming that all months of the year have the same survival probability such that $S_y = S_m^{12}$ and $\log(S_y) = 12\log(S_m)$. The relationship between the elasticities of these two survival probabilities is

$$\epsilon_{S_m} = \frac{\partial \log(\lambda)}{\partial \log(S_m)} = \frac{\partial \log(\lambda)}{\partial \log(S_y)} \frac{\partial \log(S_y)}{\partial \log(S_m)} = 12\epsilon_{S_y}$$

Elasticities to instantaneous mortality hazard rates, or their time-averaged values ($\bar{h} = -\log(S)$), are invariant to any such reparameterizations with respect to the choice of interval length of the discrete model. If this is not intuitive, consider the relationship $\epsilon_{\bar{h}} = \epsilon_{S_y} \log(S_y) = \frac{1}{12}\epsilon_{S_m} 12\log(S_m) = \epsilon_{S_m} \log(S_m)$.

1.4.1.3 Elasticity to mortality probability

As pointed out by Link & Doherty (2002), the relationship between elasticity to mortality probability, ϵ_P, and elasticity to survival probability, ϵ_S, is

$$\epsilon_P = \frac{\partial \log(\lambda)}{\partial \log(1-S)} = \frac{\partial \log(\lambda)}{\partial \log(S)} \frac{\partial \log(S)}{\partial \log(1-S)} = \epsilon_S \frac{\partial S}{\partial(1-S)} \frac{1-S}{S} = -\epsilon_S \left(\frac{1-S}{S}\right).$$

That is, a given relative decrease in morality probability only has the same impact on population growth rate (or fitness) as the same relative increase in survival probability when $S = 0.5$. Moving away from $S = 0.5$ in either direction, these two elasticities will be increasingly different. It seems hard to justify why any of these elasticities should be more relevant than the other (Link & Doherty 2002); they represent very different degrees of change in the mortality process when S is not close to 0.5. Elasticity to mortality hazard rate, on the other hand, has a clear and meaningful interpretation; it is the relative change in population growth rate caused by a relative increase in the intensity at which individuals die (Box 1 in the main text).

1.4.1.4 Elasticity to "current life-expectancy"

The relationship between the mortality hazard rate $h(t)$ and the probability density for death at time t, $f(t)$, is given as

$$h(t) = \frac{f(t)}{1 - \int_0^t f(\tau)d\tau}.$$

(Collet 2014, chap. 1.3). The probability density function $f(t)$ is here conditional on being alive at $t = 0$ (defined as time of birth or the beginning of a given interval) and integrates to 1 over infinite time. The

hazard rate function must always be greater than or equal to zero and must integrate to infinity over infinite time. It is also useful to note that the denominator in the above expression is the probability of being alive at time t, $S(t) = 1 - \int_0^t f(\tau)d\tau$. Further, the numerator of the expression is the negative derivative of this survival function. Hence, the hazard rate function can also be expressed as

$$h(t) = -\frac{\partial S(t)}{\partial t}\frac{1}{S(t)} = -\frac{\partial \log(S(t))}{\partial t}$$

(Collet 2014, chap. 1.3).

Hence, for any given probability distribution for time of death $f(t)$, there is a corresponding hazard rate function $h(t)$, and vice versa. The probability distribution for time of death corresponding to a constant hazard rate $h(t) = h$ is the exponential distribution with mean $\int_0^\infty tf(t)dt = 1/h$. Hence, if an individual (hypothetically) maintains its current mortality hazard rate throughout life, its expected time until death (i.e., its "current life-expectancy") is $1/h$. The elasticity of λ to current life-expectancy is then

$$\frac{\partial \log(\lambda)}{\partial \log(1/h)} = -\frac{\partial \log(\lambda)}{\partial \log(h)} = -\epsilon_h$$

1.4.1.5 Perspectives on hazard rates and models of stage-duration distributions

The notion of elasticity to life-expectancy introduced in section 1.4.1.4 above is similar to the notion of elasticity to stage-duration in the absence of mortality (de Valpine et al. 2014). Structured population models based on stage-duration distributions (de Valpine et al. 2014) may be specified by either specifying the stage-duration distribution itself, $f(t)$, or by specifying the corresponding hazard function $h(t)$ (see relationships in section 1.4.1.4). If individuals can leave a certain state to $K > 1$ other state ('death' being one of them), the unconditional state-duration distribution is the probability distribution corresponding to the total hazard rate $h_{tot}(t) = \sum_{k=1}^{K} h_k(t)$, where $h_k(t)$ are the transition specific hazard rates. It is, however, often of greater interest to study the probability distributions for age or time of state-transition conditional on survival. For example, Ergon et al. (2009) estimated the latent distributions[iv] for seasonal timing of onset of reproduction in populations of overwintering field voles.

Matrix population models are memory-less (Markovian) in the sense that state transition probabilities only depend on the current state of the individual. However, in individual-based "i-state configuration models" (Caswell & John 1992) one can track individual time since entering specific states, and hazard rates may in addition be dependent on "global time" and age of each individual. Such models could potentially be used to compute full conditional likelihoods (Carlin & Louis 2009), and modell fitting of capture-recapture data could be done by using Bayesian MCMC sampling. Potential applications could include estimation of (mortality) cost or reproduction depending on time since conception (or maturation), or mortality rates depending on time since infection of a pathogen.

Whenever individuals can transition from one state to more than one other state (including mortality), there is a competing risk situation among the potential transitions (Appendix S2). In addition, individuals may transition from one state to another state (or back to the original state) via other states during a given time period. A general, and very useful, tool to compute transition probability matrices from transition hazards between any number of states, involve the use of matrices of transition hazards and matrix exponentials (Keyfitz & Caswell 2005 chap. 17.5; Miller & Andersen 2008; Conn et al. 2012): First, a matrix $\mathbf{A}$ is constructed with transition hazards as the off-diagonal elements, and the negative sum of all transition hazards out of the given state as the diagonal elements, such that each row in $\mathbf{A}$ sums to zero. The corresponding transition probability matrix for an interval of length Δ is then simply

$$\mathbf{P} = \exp(\mathbf{A}\Delta).$$

[iv] By 'latent distributions' we here mean the distribution of the life-history trait that is independent of mortality, and this distribution represent the trait that is under selection. In contrast, the 'realized distribution' is the distribution of time/age of maturation among those individuals that survive until maturation. Because individuals that have a late latent maturation need to survive over a longer period before maturation can be observed, these individuals will be under-represented in the realized distribution (unless individual time/age of maturation is strongly negatively correlated with mortality hazard rate).

`exp` refers here to a matrix exponential; function 'expm' in the 'expm' package of R (Goulet et al. 2017) and function 'MatrixExponential' in the LinearAlgebra package in Maple^{TM}. See Miller & Andersen (2008) and Conn et al. (2012) for various examples. To include continuous hazard functions of time, or time since entering the states (as discussed above), one may discretize by computing transition probability matrices over many short time-intervals (e.g. days), assuming constant hazard rates within each interval.

1.4.2 Figures 5 and 6

Figures 5 and 6 illustrate the differences between elasticities of λ to S and to $\bar{h}$ by using a very simple two-stage population model without density dependence. With a post-reproductive census (Caswell 2001, p. 27), this model can be written as

$$\begin{bmatrix} N_j \\ N_a \end{bmatrix}_t = \begin{bmatrix} S_j m & S_a m \\ S_j & S_a \end{bmatrix} \begin{bmatrix} N_j \\ N_a \end{bmatrix}_{t-1}$$

where indices j and a represent respectively 'juveniles' (newborn) and 'adults' (1 year or older), N is number of individuals, S is survival probability and m is fecundity. Index value t is the time-step (year).

With a pre-reproductive census, there are no counts of newborn, and the dynamics of the population is simply described by

$$N_{a,t} = N_{a,t-1}(S_a + mS_j)$$

Hence, the population growth rate in this model is

$$\lambda = \frac{N_{a,t}}{N_{a,t-1}} = (S_a + mS_j).$$

From this we can calculate elasticities of λ to survival probabilities and their corresponding time averaged mortality hazard-rates,

$$\epsilon_{S_a} = \frac{\partial \log(\lambda)}{\partial \log(S_a)} = \frac{\partial \lambda}{\partial S_a}\frac{S_a}{\lambda} = \frac{S_a}{S_a + mS_j}$$

$$\epsilon_{S_j} = \frac{\partial \log(\lambda)}{\partial \log(S_j)} = \frac{\partial \lambda}{\partial S_j}\frac{S_j}{\lambda} = \frac{mS_j}{S_a + mS_j}$$

$$\epsilon_{\bar{h}_a} = -\bar{h}_a \epsilon_{S_a}$$

$$\epsilon_{\bar{h}_j} = -\bar{h}_j \epsilon_{S_j}$$

The elasticity ratio with respect to survival probabilities is thus

$$\frac{\epsilon_{S_a}}{\epsilon_{S_j}} = \frac{S_a}{mS_j}$$

(i.e., ϵ_{S_a} is grater that ϵ_{S_j} when $S_a > mS_j$). The elasticity ratio with respect to time-averaged mortality hazard rates is

$$\frac{\epsilon_{\bar{h}_a}}{\epsilon_{\bar{h}_j}} = \frac{\bar{h}_a}{\bar{h}_j}\frac{S_a}{mS_j}$$

Figure 5 in the main text shows the effect of reparameterizing the model with survival probabilities representing different interval lengths, and Figure 6 compares elasticity ratios with respect to survival probabilities and time-averaged mortality hazard rates. Although these figures could have been produced simply by using the relationships derived for the two-stage model above, we have chosen to do it in a more general way using the 'vitalsens' function in the 'popbio' R-package (Stubben & Milligan 2007). Hence, the code below can quite easily be adapted to any matrix population model or to other kinds or reparameterizations.

1.4.2.1 Code - Figure 5

```r
library(popbio) # For function 'vitalsens'

# Function for elastisities with reparameterization of survival probabilities
# representing different interval lengths
ela_S = function(
    Sj.y,  # Yearly juveile survival
    Sa.y,  # Yealry adult survival
    m,     # Fecundity
    t      # No. time units per year
  ){
  el = expression(  # Expressions for each of the matrix elements in a vector (by row)
    Sj.t^t*m, Sa.t^t*m,
    Sj.t^t, Sa.t^t
  )
  vr = list(Sj.t=Sj.y^(1/t), Sa.t=Sa.y^(1/t), m=m, t=t) # Num. values for demogr. params.
  v = vitalsens(el, vr)
  elas = v$elasticity
  names(elas) = rownames(v)
  return(elas)
}

# Above function allowing vector inputs for t
Ela_S = Vectorize(ela_S, "t")

# Using same structure to also reparameterize survival probabilities with respect to
# time averaged hazard rates
# This can be substantially simplified, but keeping the general structure as an example
# and for validation
ela_h = function(Sj.y, Sa.y, m, t){
  el = expression(
    exp(-hj.t)^t*m, exp(-ha.t)^t*m,
    exp(-hj.t)^t, exp(-ha.t)^t
  )
  vr = list(hj.t=-log(Sj.y^(1/t)), ha.t=-log(Sa.y^(1/t)), m=m, t=t)
  v = vitalsens(el, vr)
  elas = v$elasticity
  names(elas) = rownames(v)
  return(elas)
}

# Above function allowing vector inputs for t
Ela_h = Vectorize(ela_h, "t")

# Function for plotting
plot.panel = function(
    Sj.y, # Juvenile yearly survival probability
    Sa.y, # Adult yearly survival probability
    m,    # Fecundity
    il    # Vector of interval lengths in unit of days for reparameterization
  ){
  ES = Ela_S(Sj.y, Sa.y, m, 1/il)
  Eh = Ela_h(Sj.y, Sa.y, m, 1/il)
  plot(1,1, type="n", ylim=c(0,3), xlim=range(il), axes=F,
```

```r
       xlab = "Survival interval length",
       ylab = expression(paste("Elasticity of  ", lambda)))
  box()
  axis(2)
  axis(1, at = c(7/365, 1/4, 1/2, 1, 2), labels=c("1w", "3m", "6m", "1y", "2y"))
  abline(v = c(7/365, 1/4, 1/2, 1, 2), col="grey")
  lines(il, ES["Sj.t",], col="blue", lty=2)
  lines(il, ES["Sa.t",], col="red", lty=2)
  lines(il, ES["m",], col="green")
  lines(il, -Eh["hj.t",], col="blue")
  lines(il, -Eh["ha.t",], col="red")
  legend("topright", title = " Elasticity to ...", title.adj=0, bg="white",
         legend=c("Juvenile survival", "Adult survival", "Juvenile mortality",
                  "Adult mortality", "Fecundity"),
         col=c("blue","red","blue","red","green"), lty=c(2,2,1,1,1))
  title(paste("Juvenile survival = ", Sj.y, "\nAdult survival = ", Sa.y, sep=""))
}

# Plotting:
par(mfrow = c(1,2))
plot.panel(Sj.y = 0.2, Sa.y = 0.6, m = 2, il = (1:(2*385))/365)
title("(a)", adj=0)
plot.panel(Sj.y = 0.6, Sa.y = 0.2, m = 2, il = (1:(2*385))/365)
title("(b)", adj=0)
```

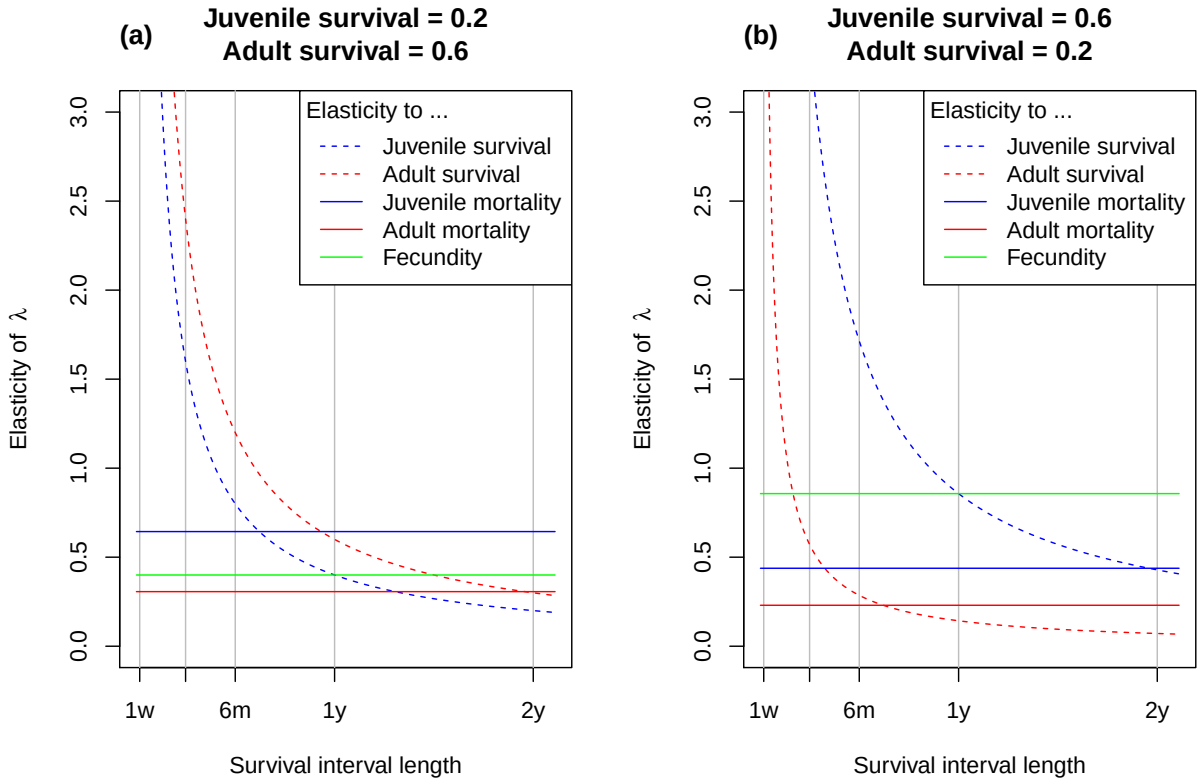
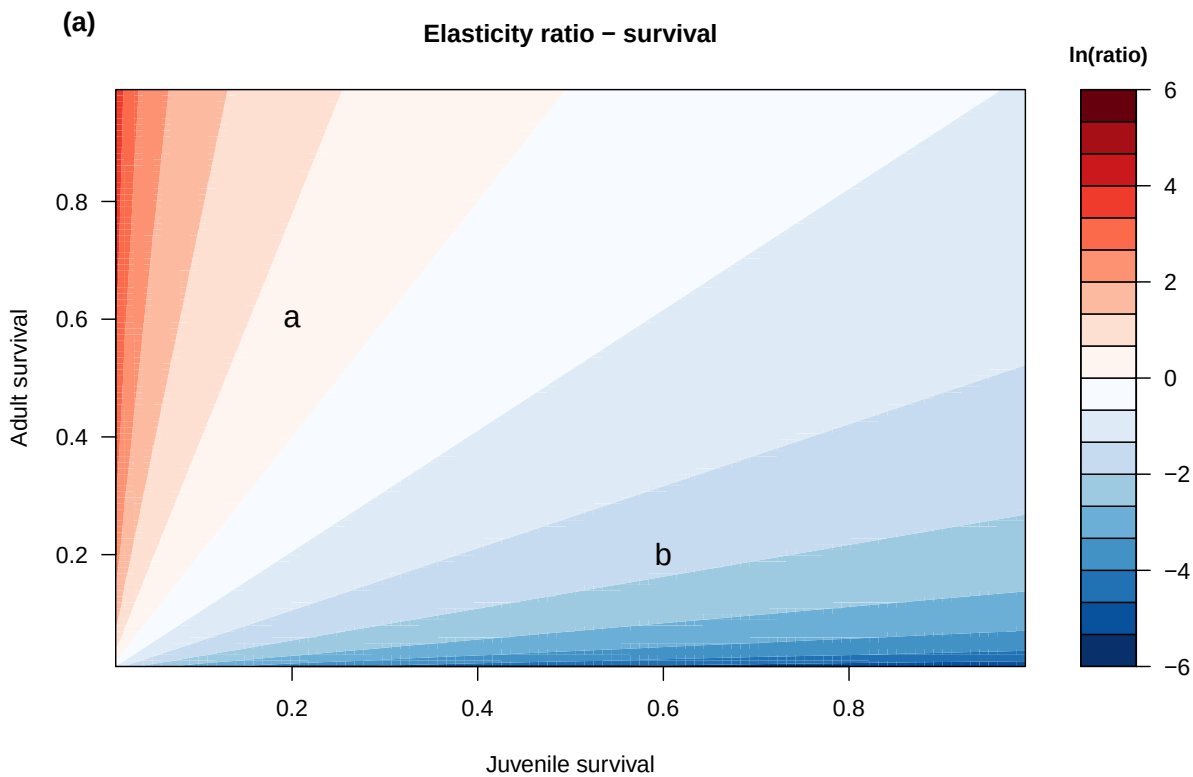


Figure 5: See caption in the main text.

1.4.2.2 Code - Figure 6

```
# Using functions from Fig. 5

library(RColorBrewer)  # for brewer.pal()

# Creating martices with elasticity ratios
n = 100
m = 2
Sj = seq(0.01,0.99,length.out=n)
Sa = seq(0.01,0.99,length.out=n)
Er.S = Er.h = matrix(NA, n, n)
for(i in 1:n){
  for(j in 1:n){
    ES = ela_S(Sj[i], Sa[j], m, 1)
    Eh = ela_h(Sj[i], Sa[j], m, 1)
    Er.S[i,j] = ES["Sa.t"]/ES["Sj.t"]
    Er.h[i,j] = Eh["ha.t"]/Eh["hj.t"]
  }
}

# Plotting
par(mfrow=c(1,2)) # Does not work with filled.contour()
Col = c(rev(brewer.pal(9,"Blues")), brewer.pal(9,"Reds"))
```

```
maxval = 6
bins = seq(-maxval, maxval, length.out=19)

filled.contour(Sj, Sa, log(Er.S), levels=bins, nlevels=18, col=Col,
        plot.axes={axis(1); axis(2); text(c(0.2, 0.6), c(0.6, 0.2), c("a","b"),
                                          cex=1.3)},
        plot.title = title("Elasticity ratio - survival", cex.main=1.1,
                                 xlab="Juvenile survival", ylab="Adult survival"),
        key.title = title(main = "ln(ratio)", cex.main=0.9, line=1))
title("(a)", outer=T, adj=0.05, line=-2)
```




```
filled.contour(Sj, Sa, log(Er.h), levels=bins, nlevels=18, col=Col,
        plot.axes={axis(1); axis(2); text(c(0.2, 0.6), c(0.6, 0.2), c("a","b"), cex=1.3)},
        plot.title = title("Elasticity ratio -  - mortality rate", cex.main=1.1,
                                 xlab="Juvenile survival", ylab="Adult survival"),
        key.title = title(main = "ln(ratio)", cex.main=0.9, line=1))
title("(b)", outer=T, adj=0.05, line=-2)
```

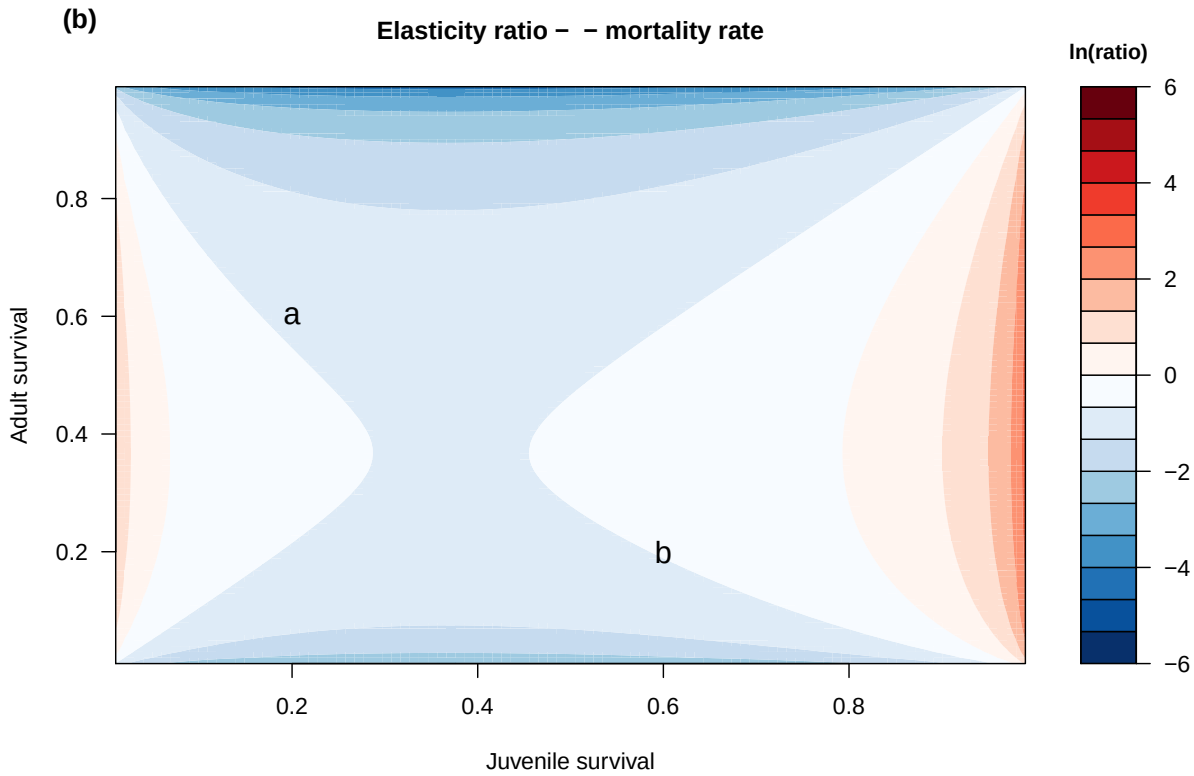


Figure 6: See caption in the main text.

1.4.3 Figure 7

As shown in Figure 6 above, patterns in elasticities to mortality hazard rate are very different from patterns in elasticities to survival probability. One way to understand this is to note, as shown in Figure 7, that the relationship between $\log(S)$ and $\log(h)$ is very non-linear ($\log(S) = -\exp(\log(h))$ (the first derivative in this plot is $-h$). When the hazard rate is low (survival is high), a large relative change in the hazard rate (to the left in the figure) corresponds to a small relative change in survival probability. When the hazard rate is high (survival is low), a small relative change in the hazard rate corresponds to a large relative change in survival. This is also reflected in the relationship between elasticities to survival probability and elasticity to mortality hazard rate given above; $\epsilon_{\bar{h}} = -h\epsilon_S$.

1.4.3.1 Code - Figure 7

```
S = seq(0.001, 0.999, length.out=40)
y = log(S)
x = log(-log(S))

par(mai=c(1.2,1.2,1.2,1.2))
plot(x,y, xlab="ln(Mortality hazard rate)", ylab="ln(Survival probability)", type="n")

lS = c(0.001, 0.01, 0.1, 0.5, 0.9)

axis(3, at = log(-log(lS)), labels=as.character(lS), col.ticks="darkgrey",
```

```
     col.axis="darkgrey")
axis(4, at = log(lS), labels=as.character(lS), col.ticks="darkgrey", col.axis="darkgrey")
abline(h=log(lS), v=log(-log(lS)), col="darkgrey")
mtext("Survival probability",3, line=2.5, col="darkgrey")
mtext("Survival probability",4, line=2.5, col="darkgrey")
lines(x,y, lwd=2)
```

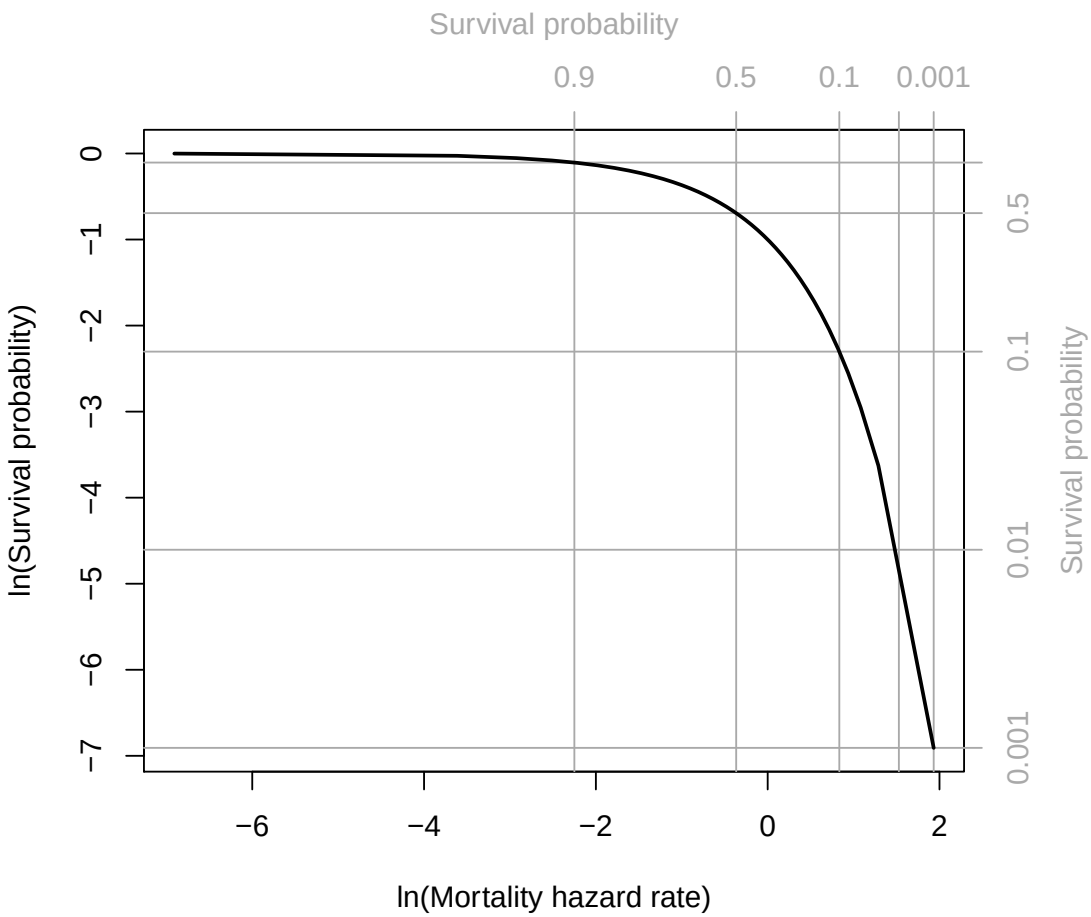


Figure 7: See caption in the main text.

1.4.4 Variance-stabilized sensitivities

In the final paragraph of this section in the main text, we comment on the use of variance-stabilized sensitivities (VSS) as defined by Link & Doherty (2002). VSS are based on an assumed relationship between the temporal variance of a demographic parameter θ and the temporal mean μ of this parameter, which may be described as $\mathrm{Var}(\theta) = f(\mu)$. The delta-method approximation to the variance of a transformation $q(\theta)$ of this parameter

is then

$$\mathrm{Var}(q(\theta)) \approx \left(\frac{\partial q(\mu)}{\partial \mu}\right)^2 \mathrm{Var}(\theta) = \left(q'(\mu)\right)^2 f(\mu).$$

For a given variance-mean relationship $f(\mu)$ the transformation $q(\theta)$ is then said to be a variance-stabilizing transformation when $\mathrm{Var}(q(\theta))$ is independent of μ, i.e., when $\left(q'(\mu)\right)^2 f(\mu)$ is a constant. Link & Doherty (2002) then defined the variance-stabilized sensitivity for a given variance-mean relationship as the sensitivity of $\log(\lambda)$ to the transformed parameter $q(\theta)$

$$\mathrm{VSS}_q(\lambda, \theta) = \frac{\partial \log(\lambda)}{\partial q(\theta)} = \frac{\partial \lambda}{\partial \theta} \frac{\partial \theta}{\partial q(\theta)} \frac{1}{\lambda} = \frac{\partial \lambda}{\partial \theta} \frac{1}{\lambda q'(\theta)}$$

To find the variance-mean relationship that corresponds with a particular variance stabilizing transformation $q(\theta)$, we need to solve $\left(q'(\mu)\right)^2 f(\mu) = C$ for $f(\mu)$ where C is an arbitrary positive constant. When $q(\theta)$ is a logit-transformation of survival probability, we have $q(S) = \log\left(\frac{S}{1-S}\right)$ and hence

$$f(\mu) = C\left(q'(\mu)\right)^{-2} = C\left(\frac{1}{\mu(1-\mu)}\right)^{-2} = C\left(\mu(1-\mu)\right)^2$$

Similarly, the loglog-link on survival probability, $q(S) = \log(-\log(S))$, is a variance stabilizing transformation when

$$f(\mu) = C\left(q'(\mu)\right)^{-2} = C\left(\frac{1}{\mu \log(\mu)}\right)^{-2} = C\left(\mu \log(\mu)\right)^2$$

When using the loglog-link on survival probability, $q(S) = \log(-\log(S)) = \log(\bar{h})$, as a variance stabilizing transformation, the variance-stabilized sensitivity becomes

$$\mathrm{VSS}_q(\lambda, S) = \frac{\partial \log(\lambda)}{\partial q(\theta)} = \frac{\partial \log(\lambda)}{\partial \log(\bar{h})} = \epsilon_{\bar{h}}$$

Hence, using the loglog-link to obtain a variance-stabilized sensitivity to survival probability, assuming that the temporal variance in interval specific survival probabilities is proportional to $\left(\mu \log(\mu)\right)^2$, is equivalent to using the elasticity to time-averaged mortality hazard rate.

Appendix S2: Competing risks and cause-specific mortality probabilities

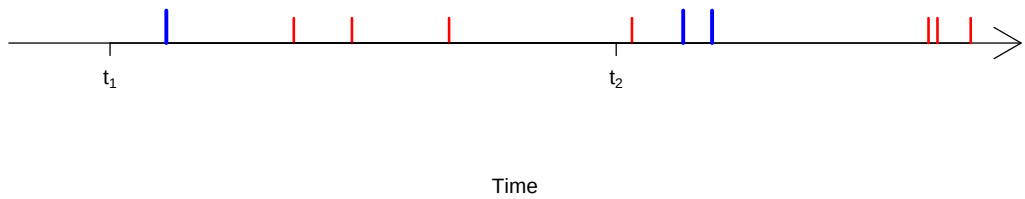


Figure S2.1: Realized times of two types of independent recurrent events (or "deadly events experienced by an immortal individual")

Figure S2.1 illustrates a situation where another ("blue") cause of mortality has been added to the "red" cause of mortality illustrated in the figure in Box 1 in the main text. If we, for simplicity, assume that the hazard rates for these two causes of mortality are constant over an interval with values h_1 and h_2, the total mortality hazard rate is now $h_1 + h_2$, and the probability of surviving the interval of length Δ is $S = e^{-(h_1+h_2)\Delta}$.

It is essential to realize that adding the blue cause of mortality in this example reduces the probability of dying from the red cause, even though the hazard rate representing the red cause remains the same. This is due to the fact that the individual now may die from a blue cause before a red event takes place. The probability of dying from a specific cause (often referred to as a "risk"), when the hazard rate of this cause remains unchanged, will always go down when the hazard rate of other mortality causes increase (or go up if the hazard rate of other causes decrease). One may say that the risks of death are "competing", and the probability of "winning" depends on the strength of the competitors (i.e., mortality probabilities are intrinsically linked). Hence, a change in the probability of dying of a specific cause may be due to changes in any of the cause specific mortality hazard rates. When studying e.g., environmental influence on different causes of mortality, or how different causes of mortality vary and co-vary among groups of individuals or among years within a population, it is therefore of prime interest to estimate the hazard rates of different causes of mortality rather than the probabilities of dying from different causes.

Discrete-time models (such as classical capture recapture models) that incorporate different causes of mortality involve state-transition probability parameters representing the probabilities that an individual that is alive at one time-point (t_1) is either alive or dead from specific causes at a later point in time (t_2). These transition probabilities should sum to 1. Hence, $S + \sum_{k=1}^{K} P_k = 1$, where S is the survival probability and the P_k's are the probabilities of dying from different causes classified in K categories. Assuming, for simplicity, that each of the cause specific hazard rates, h_k are constant within intervals, we know from above that $S = e^{-\sum_{k=1}^{K} h_k \Delta}$. To find the probability of dying from cause k, P_k, we may use the definition of hazard rates as

$$h(t) = \lim_{dt \to 0} \Pr(t \leqslant T < t + dt | t < T)/dt$$

(Collett 2014) where T is time of death and dt is an infinitesimally small unit of time (i.e., $h(t)dt$ is the probability that the individual will die before time $t + dt$ given that it is alive at time t). From this we see

that the relationship between the hazard rate and the probability density for time of death at time t, $f(t)$, is

$$h(t) = \frac{f(t)}{S(t)}$$

where $S(t)$ is the probability of being alive at time t, which equals $1 - \int_0^t f(t)dt$. Using this relationship, we can now find the probability density function for the time of death from cause k, $f_k(\tau)$, where $\tau = t - t_1$ is the time since the beginning of the interval,

$$f_k(\tau) = h_k e^{-\tau \sum_{l=0}^{K} h_l}$$

Note that this density function is not a proper probability distributions function as it does not integrate to 1 over infinite time. Instead, it integrates to the probability of ever dying from cause k, which in the case of constant hazard rates becomes $h_k / \sum_{l=1}^{K} h_l$. To find the probability of dying from cause k before the end of the interval, we need to integrate $f_k(\tau)$ from $\tau = 0$ to $\tau = \Delta$ $(= t_2 - t_1)$, and get

$$P_k = \int_0^{\Delta} f_k(\tau)d\tau = \left(1 - e^{-\Delta \sum_{l=1}^{K} h_l}\right) \frac{h_k}{\sum_{l=1}^{K} h_l} = (1 - S) \frac{h_k}{\sum_{l=1}^{K} h_l}$$

This expression can be generalized to cases where cause specific hazard rates may vary over time but remain proportional throughout the interval such that $\alpha_k = h_k(t) / \sum_{l=1}^{K} h_l(t)$, in which case we get

$$P_k = (1 - S)\alpha_k$$

Note that the conditional probability of having died from cause k, given that the individual has died, in the case of proportional hazard rates is the fraction of the total mortality hazard rate attributed to the specific cause, α_k, which in this case equals the probability of ever dying of this cause. This is, however, not the case when hazard rates are not proportional over time.

Section 1.4.1.5 in Appendix S1 outlines a more general approach to deriving/computing transition probabilities under competing risk.

Appendix S3: Fitting capture-recapture models with cause-specific mortality

3.1 Introduction

When information about cause of death (e.g., harvesting) is available for at least a subset of marked individuals in the population, multi-state capture-recapture models can be fitted to the data to address questions relating to e.g. compensatory effects of harvesting or age-related patterns in mortality from different causes. However, commonly used software for fitting multi-state capture-recapture data such as MARK (White & Burnham 1999) and E-SURGE (Choquet, Rouan & Pradel 2009) use the multinomial logit-link function to constrain rows of the transition matrix to sum to 1. For example, the probability of transitioning from the alive state to death-by-cause state k (i.e., probability of dying from cause k) may be modelled as

$$P_k = \frac{\exp(\eta_k)}{1 + \sum_{l=1}^{K} \exp(\eta_l)},$$

while the probability of surviving is

$$S = 1 - \sum_{k=1}^{K} P_k = \frac{1}{1 + \sum_{l=1}^{K} \exp(\eta_l)},$$

where η_k is the linear predictor relating to cause k. This multinomial link-function does not facilitate modelling of cause-specific hazard rates (Appendix S2). Instead, based on the expressions for S and P_k derived in Appendix S2, we can model log-linear effects on the mortality hazard rates by using the multinomial link-function defined by

$$S = \exp\left(-\sum_{k=1}^{K} \exp(\eta_k)\right)$$

and

$$P_k = (1 - S)\,\frac{\exp(\eta_k)}{\sum_{l=1}^{K} \exp(\eta_l)}$$

To fit, for example, a model with two causes of mortality one needs a 4 by 4 transition matrix

$$\mathbf{\Psi} = \begin{bmatrix} S & P_1 & P_2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The four states are here 'alive', 'newly dead from cause 1', 'newly dead from cause 2' and a final absorbing state (rows represent "from-state" and columns represent "to-state"). The distinction between newly dead states and the final state is necessary because individuals must be reported dead in the year they died (i.e., year of death of recovered individuals must be known). Recapture/recovery probabilities of the four states are

$$\mathbf{p} = \begin{bmatrix} p_0 & p_1 & p_2 & 0 \end{bmatrix}$$

In the most general (full, unconstrained) model, parameters would be individual and year specific, but indices are omitted here for simplicity. Survival probabilities in the general multi-state capture-recapture models must be fixed to 1 for the three first states and to 0 (or any probability value) for the absorbing state.

To facilitate modelling of mortality hazard rates, we modified the functions used for fitting multi-state capture-recapture models in the R-package 'marked' package (Laake, Johnson & Conn 2013). This package combines a system for highly flexible model formulations (including modelling of covariate effects relating to time, age, cohort and individuals), similar to the popular 'RMark' package, with highly efficient numerical optimization of the likelihood function using automatic differentiation through executables compiled by

the ADMB software (Fournier et al. 2012). The files are available on https://github.com/torbjore/study_mortality_with_hazard_rates, and a worked example using simulated data is included below.

Schaub & Pradel (2004) and Schaub (2009) noted that only some sub-models of another version of the general multi-state capture-recapture model (see Table 1 in the main text) are identifiable. This is also the case for our model, but a detailed treatment of identifiability is beyond our scope here.

3.2 Preparation

To use the modified 'marked' files, an executable must first be compiled and saved on your computer. To do this, first install the ADMB software from http://www.admb-project.org/. ADMB needs a C++ compiler, but this is included in some of the ADMB installation files. It is safest to install ADMB in a directory without spaces in its name and path. The code below assumes that ADMB is installed in "C:/admb", but this can be modified. Once ADMB and a C++ compiler is installed, the following code will create a 'multistate_csmhr.exe' file in your working directory (use install.packages() to install any missing packages).

```
library(marked)
library(R2admb)
library(RCurl) # For getURL

prepare_admb = function()
{
  Sys.setenv(PATH = paste("c:/admb/bin;c:admb/utilities;c:/admb/utilities/mingw64/bin;",
                          Sys.getenv("PATH"), sep = ";"))
  Sys.setenv(ADMB_HOME = "c:/admb")
  invisible()
}
prepare_admb()

# Downloading the ADMB tpl file and building a local executable
# Breaking up string to fit in box
reprstr1 = "https://raw.githubusercontent.com/torbjore/study_mortality_with_hazard_rates/"
reprstr2 = "master/modified_marked_files/"
tplfile = paste(reprstr1, reprstr2, "multistate_csmhr.tpl", sep="")
tpl = getURL(tplfile)
write(tpl, "multistate_csmhr.tpl")
marked::setup_admb("multistate_csmhr", compile=TRUE, clean=FALSE)
```

3.3 Example

In this simple example we will simulate 10 years of capture-recapture data with two causes of mortality; 'hunting' and 'natural'. We will assume that both mortality hazard rates varies independently over time, that probability of recapture of an individual alive is 0.5, that the reporting probability of a hunted individual is 0.7, and that the probability that an individual which died of other (natural) causes will be found and reported is 0.1.

First, we simulate data:

```
set.seed(11) # for reproducibility

p = c(0.5, 0.7, 0.1, 0)
N.ind = 200 # Number of individuals marked each year
N.yr = 10   # Number of years
```

```r
# Year-specific hazard rates for the two causes
h1 = runif(N.yr-1, 1/25, 1/2)
h2 = runif(N.yr-1, 1/25, 1/2)

# For computation of transition matrix
fTM = function(h1, h2){
  # Yearly survival probability
  S = exp(-(h1+h2))
  # Yearly mortality probabilities
  P1 = (1-S)*h1/(h1+h2)
  P2 = (1-S)*h2/(h1+h2)
  # Transition matrix
  tm = matrix(
    c(
      S, P1, P2, 0,
      0, 0, 0, 1,
      0, 0, 0, 1,
      0, 0, 0, 1
    ), 4, 4, byrow = T
  )
  return(tm)
}

first = matrix(rep(1:(N.yr - 1), each = N.ind), ncol=1) # When marked

# Simulating Matrix of true states
N.ind.tot = length(first)
St = matrix(NA, N.ind.tot, N.yr)
for(i in 1:N.ind.tot){
  St[i,first[i]] = 1 # All individuals introduced as alive
  for(j in first[i]:(N.yr-1)){
    St[i,j+1] = sample(1:4, 1, prob = fTM(h1[j], h2[j])[St[i,j],])
  }
}
St[is.na(St)] = 4 # to get zero capture probability before marking

# Computing matrix of state-dependent capture probabilities
P = matrix(p[St], N.ind.tot, N.yr)

# Simulating Capture History data
capt = rbinom(N.ind.tot*N.yr,1,P)
CH = St*capt
for(i in 1:N.ind.tot){
  CH[i,first[i]] = 1 # Conditioning on first capture; hence inds are always seen when marked
}
```

Then we source in the modified `marked` functions and prepare the data for model fitting:

```r
library(marked)
library(R2admb)
prepare_admb() # Function defined above

# Source in modified functions from Git-hub
crm_csmhr_FUN = paste(reprstr1, reprstr2, "crm_csmhr.R", sep="")
```

```r
mscjs_csmhr_FUN = paste(reprstr1, reprstr2, "mscjs_csmhr.R", sep="")
source(crm_csmhr_FUN)
source(mscjs_csmhr_FUN)

# Transforming data to a sting variable for 'marked'
simdata = data.frame(
  ch = apply(CH, 1, function(i) paste(c("0","A","B","C")[i+1], collapse = ""))
)
simdata$ch = as.character(simdata$ch)

# Adding an individual seen in state D at the *last* occasion to get the right dimension
# of the transitioin matrix. Since this individual is captured only in the last year,
# and we condition on first capture, this individual will not affect estimates (and will
# be removed internally)

simdata = rbind(simdata, ch = paste(paste(rep("0", N.yr-1), collapse = ""), "D", sep=""))

simdata.processed = process.data(simdata,model="Mscjs",strata.labels=c("A","B","C","D"))
simdata.ddl = make.design.data(simdata.processed)
```

The resultant `simdata.dll` object is a list of data frames with predictor variables for all parameters in the full (unconstrained) model; one data.frame for modelling of log hazard rates (`simdata.ddl$Psi`), and one data.frame for modelling of logit recapture/recovery probabilities (`simdata.ddl$p`). Individual covariates, grouping variables and individual ages at first capture may be included in `process.data()` (see `help(package="marked")` for documentation).

Below, we create variables `simdata.ddl$Psi$hunting` and `simdata.ddl$Psi$natural`, use this to constrain the model and fit it by using the modified functions:

```r
simdata.ddl$Psi$hunting = ifelse(simdata.ddl$Psi$tostratum == "B", 1, 0)
simdata.ddl$Psi$natural = ifelse(simdata.ddl$Psi$tostratum == "C", 1, 0)

# Submodel constraints
model = list(
  p = list(formula = ~ -1 + stratum),
  Psi = list(formula = ~ -1 + natural:time + hunting:time)
)

# NB! Convergence may be sensitive to starting values. It's a good idea to try many
# different starting values to check for convergence at local maxima:
inits = list(p=rep(0,3), Psi = rep(mean(log(c(h1,h2))), 2*(N.yr-1)))

mod1_csmhr = crm_csmhr(simdata.processed, simdata.ddl, clean=F, model.parameters=model,
                       initial=inits, hessian=TRUE)
```

```
##
## Elapsed time in minutes:  0.1292
```

```r
# NB! Run with clean=F, or else the executable will be deleted!
```

Note that is not necessary to fix the survival probabilities or fixed transition matrix elements as this is done internally in the `crm_csmhr` function.

To see parameter estimates and a summary of the model fit, type the name of the model object

```r
mod1_csmhr
```

```
##
## crm Model Summary
##
## Npar :  21
## -2lnL:  8687.64
## AIC  :  8729.64
##
## Beta
##                      Estimate         se        lcl         ucl
## p.stratumA        -0.04400161 0.05217081 -0.1462564  0.05825318
## p.stratumB         0.54749248 0.39423422 -0.2252066  1.32019156
## p.stratumC        -2.21528017 0.17240226 -2.5531886 -1.87737173
## Psi.natural:time1 -2.13366000 0.53993877 -3.1919400 -1.07538000
## Psi.natural:time2 -2.54077677 0.38698251 -3.2992625 -1.78229105
## Psi.natural:time3 -1.76863318 0.29324920 -2.3434016 -1.19386475
## Psi.natural:time4 -0.63110307 0.10472320 -0.8363605 -0.42584559
## Psi.natural:time5 -0.97332549 0.13576997 -1.2394346 -0.70721634
## Psi.natural:time6 -1.16448218 0.25093775 -1.6563202 -0.67264419
## Psi.natural:time7 -1.25585599 0.17927273 -1.6072305 -0.90448144
## Psi.natural:time8 -1.38657207 0.22280072 -1.8232615 -0.94988265
## Psi.natural:time9 -2.02119027 0.47154286 -2.9454143 -1.09696626
## Psi.time1:hunting -1.32971130 0.24566097 -1.8112068 -0.84821580
## Psi.time2:hunting -2.93815957 0.32592799 -3.5769784 -2.29934071
## Psi.time3:hunting -1.20405839 0.17919551 -1.5552816 -0.85283519
## Psi.time4:hunting -2.85968339 0.30819935 -3.4637541 -2.25561266
## Psi.time5:hunting -2.49942724 0.26982471 -3.0282837 -1.97057081
## Psi.time6:hunting -0.77625675 0.17492928 -1.1191181 -0.43339537
## Psi.time7:hunting -2.15652028 0.24773121 -2.6420735 -1.67096710
## Psi.time8:hunting -1.65644715 0.20087128 -2.0501549 -1.26273944
## Psi.time9:hunting -0.66242305 0.17041356 -0.9964336 -0.32841248
```

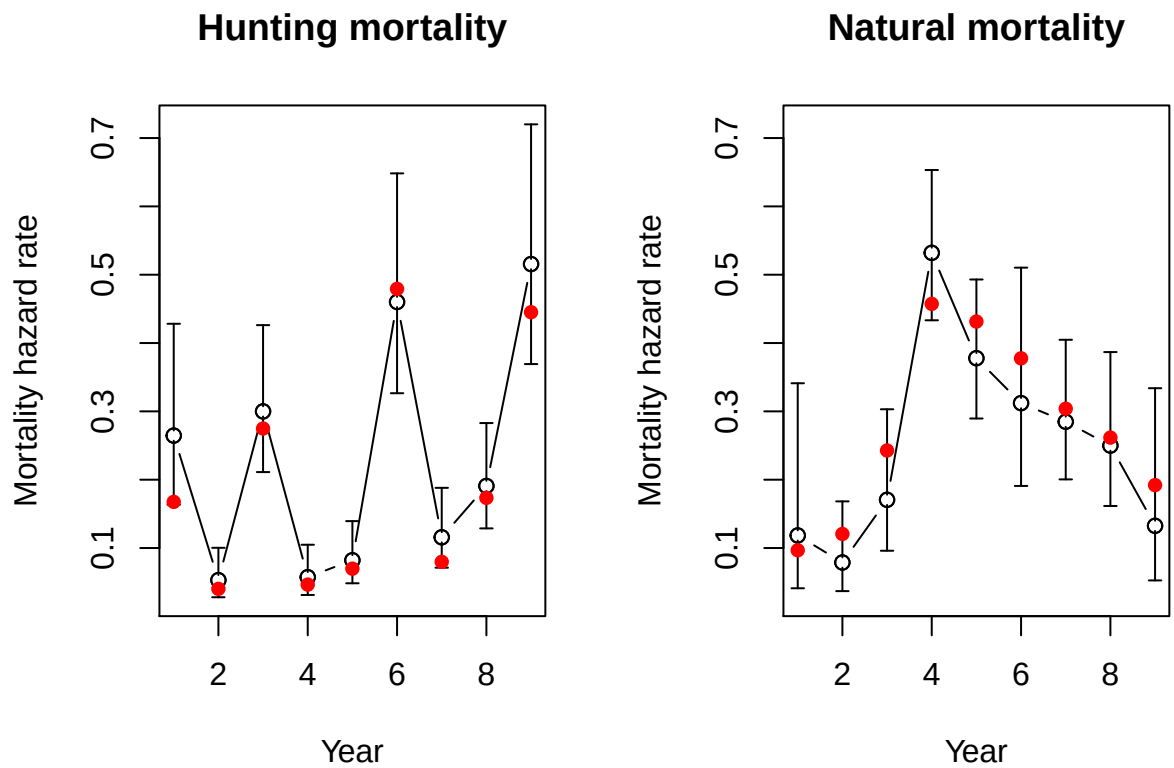
Here, `p` parameters are logit-linear coefficients in the model for recapture/recovery probabilities, and `Psi` parameters are coefficients in the log-linear models for each of the mortality hazard rates.

Please note that convergence on a local likelihood peak is not unlikely, and it is highly recommendable to try many different starting values (lower AIC values indicate higher likelihood values). It is also possible to change convergence criteria and perform various diagnostics (as well as computing likelihood profiles, confidence intervals by MCMC, etc) by passing ADMB command line options through the `extra.args` argument in `crm_csmhr` or by modifying the `multistate_csmhr.tpl` file - see the ADMB manual at http://www.admb-project.org/.

The `marked::predict.crm` function cannot be used for prediction of mortality hazard rates (only for recapture/recovery probabilities), but we include a simple function for prediction below. Here we compute and plot the predicted mortality hazard rates and compare them to the true values used in the simulation:

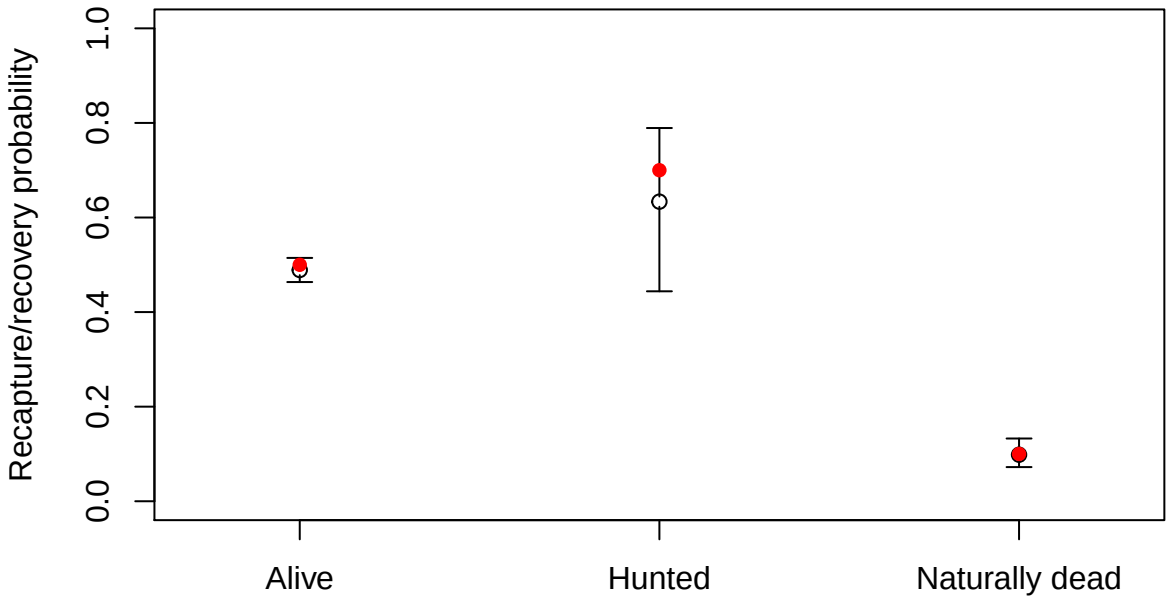
```r
predict.marked.csmhr = function(fit, newdata, par, invlink){
  model = fit$model.parameters[[par]]$formula
  X = model.matrix(model, newdata)
  b = fit$results$beta[[par]]
  vc = fit$results$beta.vcv
  nms = dimnames(vc)[[1]]
  use = substring(nms,1,nchar(par)) == par
  vc = vc[use,use]
```

```r
  eta = X %*% b
  se = sqrt(diag(X %*% vc %*% t(X)))
  D = newdata
  D$eta = eta
  D$se = se
  D$est = invlink(eta)
  D$lwr = invlink(eta - 1.96*se)
  D$upr = invlink(eta + 1.96*se)
  return(D)
}

D = expand.grid(year = 1:9, cause = c("hunting","natural"))
D$time = factor(paste("time", D$year, sep=""))
D$hunting = ifelse(D$cause == "hunting", 1, 0)
D$natural = ifelse(D$cause == "natural", 1, 0)

Pred = predict.marked.csmhr(fit = mod1_csmhr, newdata = D, par = "Psi",
                            invlink = function(x) exp(x))

# Adding true values
Pred$true = ifelse(Pred$hunting==1, h1, h2)

# Plotting
library(plotrix) # for plotCI

par(mfrow=c(1,2))
Ylim = range(c(Pred$lwr, Pred$upr))
with(Pred[Pred$cause=="hunting",],
     plotCI(year, est, ui=upr, li=lwr, gap=T, ylim=Ylim,
            xlab = "Year", ylab = "Mortality hazard rate"))
with(Pred[Pred$cause=="hunting",], lines(year, est, type="b"))
with(Pred[Pred$cause=="hunting",], points(year,true, pch=16, col="red"))
title("Hunting mortality")

with(Pred[Pred$cause=="natural",],
     plotCI(year, est, ui=upr, li=lwr, gap=T, ylim=Ylim,
            xlab = "Year", ylab = "Mortality hazard rate"))
with(Pred[Pred$cause=="natural",], lines(year, est, type="b"))
with(Pred[Pred$cause=="natural",], points(year,true, pch=16, col="red"))
title("Natural mortality")
```



We do the same for recapture/recovery probabilities:

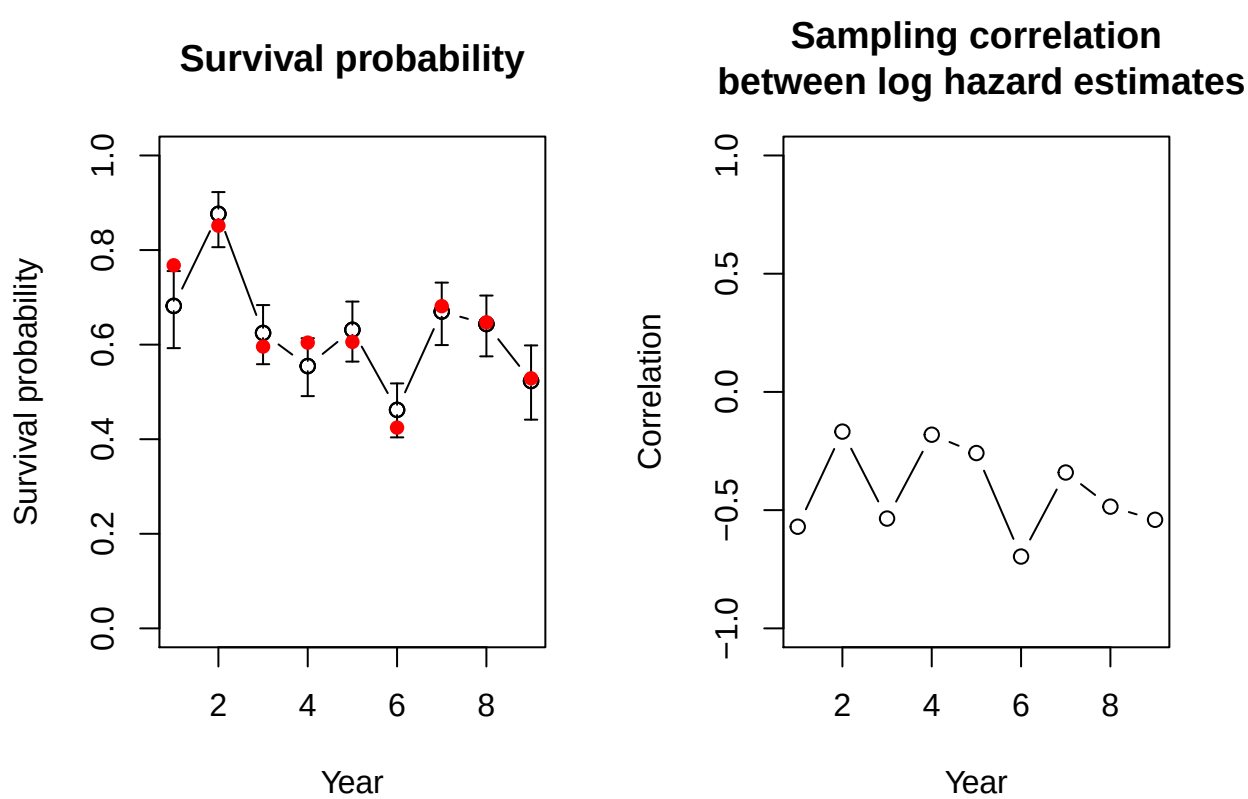
```
# Recapture/recovery probability
D.p = data.frame(stratum = c("A","B","C"))
Pred.p = predict.marked.csmhr(fit = mod1_csmhr, newdata = D.p, par = "p",
                              invlink = function(x) exp(x)/(1+exp(x)))
Pred.p$true = p[1:3]

Ylim = c(0,1)
with(Pred.p,
     plot(1:3, est, ylim=Ylim, xlim=c(0.7,3.3),
          xlab = "", ylab = "Recapture/recovery probability", type="n", axes=F))
with(Pred.p, plotCI(1:3, est, add=T, ui=upr, li=lwr, gap=T))
with(Pred.p, points(1:3, true, pch=16, col="red"))
box()
axis(2)
axis(1, at=1:3, labels = c("Alive", "Hunted", "Naturally dead"))
```



Finally, we use the delta-method to compute approximate standard errors of the log of total mortality hazard rates and plot the confidence intervals of survival probabilities. At the same time, we compute and plot estimates of the sampling correlation between the two log hazard rates for each year (may indicate identifiability problems if highly negative).

```
# Function for a single year
S.delta = function(yr, fit, D){
  DD = D[D$year==yr,]
  model = fit$model.parameters[["Psi"]]$formula
  X = model.matrix(model, DD)
  b = fit$results$beta[["Psi"]]
  vc = fit$results$beta.vcv
  nms = dimnames(vc)[[1]]
  use = substring(nms,1,3) == "Psi"
  vc = vc[use,use]
  log.h = X %*% b
  VC.log.h = X %*% vc %*% t(X)
  h = exp(log.h)
  h.tot = sum(h)
  log.h.tot = log(h.tot)
  x = h/h.tot
  var.log.h.tot = t(x) %*% VC.log.h %*% x
  se.log.h.tot = sqrt(diag(var.log.h.tot))
  lwr.log.h.tot = log.h.tot - 2*se.log.h.tot
  upr.log.h.tot = log.h.tot + 2*se.log.h.tot
  list(
    S = exp(-h.tot),
```

```
    lwr.S = exp(-exp(upr.log.h.tot)),
    upr.S = exp(-exp(lwr.log.h.tot)),
    cor.log.h = cov2cor(VC.log.h)[1,2]
  )
}

S.est = sapply(1:(N.yr-1), S.delta, fit=mod1_csmhr, D=D)
S.est = as.data.frame(t(S.est))

par(mfrow=c(1,2))
plot(1:(N.yr-1), unlist(S.est$S), ylim=c(0,1),
     xlab = "Year", ylab = "Survival probability")
with(S.est,
     plotCI(1:(N.yr-1), unlist(S), ui=unlist(upr.S), li=unlist(lwr.S), add=T, gap=T))
lines(1:(N.yr-1), unlist(S.est$S), type="b")
points(1:(N.yr-1),exp(-(h1+h2)), pch=16, col="red")
title("Survival probability")

plot(1:(N.yr-1), unlist(S.est$cor.log.h), ylim=c(-1,1),
     xlab = "Year", ylab = "Correlation", type="b")
title("Sampling correlation\n between log hazard estimates")
```



References

Cam, E., Link, W.A., Cooch, E.G., Monnat, J.Y. & Danchin, E. (2002) Individual covariation in life-history traits: Seeing the trees despite the forest. American Naturalist, 159, 96-105.

Carlin, B.P. & Louis, T.A. (2009) Bayesian Methods for Data Analysis. Chapman & Hall.

Caswell, H. & John, A.M. (1992) From the individual to the population in demographic models. Individual-based models and approaches in ecology (eds D.L. DeAngelis & L.J. Gross). Chapman & Hall, New York, NY.

Choquet, R., Rouan, L. & Pradel, R. (2009) Program E-SURGE: a software application for fitting multievent models. Modeling Demographic Processes in Marked Populations (eds D.L. Thompson, E.G. Cooch & M.J. Conroy). Springer.

Collett, D. (2014) Modelling survival data in medical research, Third edn. CRC Press.

de Valpine, P., Scranton, K., Knape, J., Ram, K. & Mills, N.J. (2014) The importance of individual developmental variation in stage-structured population models. Ecol Lett, 17, 1026-1038.

Ergon, T. & Gardner, B. (2014) Separating mortality and emigration: modelling space use, dispersal and survival with robust-design spatial capture–recapture data. Methods in Ecology and Evolution, 5, 1327-1336.

Ergon, T., Yoccoz, N. & Nichols, J.D. (2009) Estimating latent time of maturation and survival costs of reproduction in continuous time from capture-recapture data. Modeling Demographic Processes in Marked Populations (eds D. Thomson, E.G. Cooch & M.J. Conroy). Springer Science+Business Media, LLC.

Fournier, D.A., Skaug, H.J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M.N., Nielsen, A. & Sibert, J. (2012) AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. Optim. Methods Softw., 27, 233-249.

Goulet, V., Dutang, C., Maechler, M., Firth, D., Shapira, M. & Stadelmann, M. (2017). expm: Matrix Exponential, Log, 'etc'. R package version 0.999-2. <https://CRAN.R-project.org/package=expm>.

Laake, J.L., Johnson, D.S. & Conn, P.B. (2013) marked: An R package for maximum-likelihood and MCMC analysis of capture-recapture data. Methods in Ecology and Evolution, 4, 885-890.

Langtimm, C.A. (2009) Non-random temporary emigration and the robust-design: conditions for bias at the end of a time series. Modelling Demographic Processes in Marked Populations (eds D.L. Thomson, E.G. Cooch & M.J. Conroy). Springer Science+Business Media, LLC.

Link, W.A. & Doherty, P.F. (2002) Scaling in sensitivity analysis. Ecology, 83, 3299-3305.

Miller, T.J. & Andersen, P.K. (2008) A finite-state continuous-time approach for inferring regional migration and mortality rates from archival tagging and conventional tag-recovery experiments. Biometrics, 64, 1196-1206.

Quinn, T.J. & Deriso, R.B. (1999) Quantitative fish dynamics. Oxford University Press.

Ricker (1975) Computation and interpretation of biological statistics of fish populations. Bull. Fish. Res. Board Can., 191.

Schaub, M. (2009) Evaluation of bias, precision and accuracy of mortality cause proportion estimators from ring recovery data. Modeling Demographic Processes in Marked Populations (eds D. Thomson, E.G. Cooch & M.J. Conroy). Springer Science+Business Media, LLC.

Schaub, M. & Pradel, R. (2004) Assessing the relative importance of different sources of mortality from recoveries of marked animals. Ecology, 85, 930.

Stubben, C.J. & Milligan, B.G. (2007) Estimating and Analyzing Demographic Models Using the popbio Package in R. Journal of Statistical Software, 22(11).

Vaupel, J.W., Manten, K.G. & Stallard, E. (1979) The impact of heterogeneity in individual frailty on the dynamics of mortality. Demography, 16, 439-454.

White, G.C. & Burnham, K.P. (1999) Program MARK: Survival estimation from populations of marked animals. Bird Study, 46 Supplement, 120-138.